

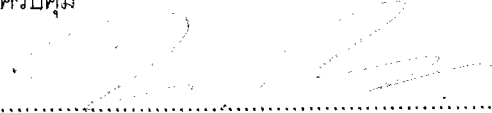
การแสดงความเที่ยงตรงและความเชื่อมั่นของแบบทดสอบปลายเปิด วิชาคณิตศาสตร์

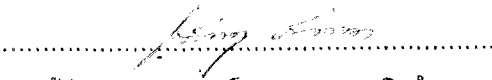
ปริญาพนธ์
ของ
อรอนงค์ บำรุง

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2542
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

คณะกรรมการควบคุมและคณะกรรมการสอบได้พิจารณาปริญญานิพนธ์ฉบับนี้แล้ว
เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร ปริญญาการศึกษามหาบัณฑิต
วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒได้

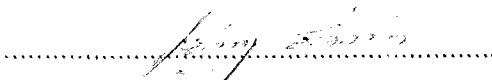
คณะกรรมการควบคุม

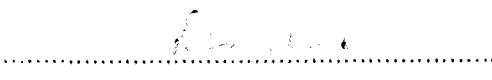

..... ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์)

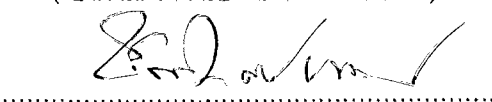

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ เขาวนา ขวลิขิตารัง)

คณะกรรมการสอบ

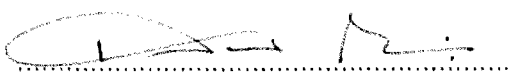

..... ประธาน
(รองศาสตราจารย์ ดร.บุญเชิด ภิญโญอนันตพงษ์)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์ เขาวนา ขวลิขิตารัง)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(รองศาสตราจารย์ นิภา ศรีไพโรจน์)


..... กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ ขวลิขิต รวยอาจิน)

บัณฑิตวิทยาลัยอนุมัติให้ปริญญานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร ปริญญา
การศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา ของมหาวิทยาลัยศรีนครินทรวิโรฒ


..... คณบดีบัณฑิตวิทยาลัย
(ศาสตราจารย์ ดร. เสริมศักดิ์ วิศาลาภรณ์)

วันที่ 14 เดือน พฤษภาคม พ.ศ. 2542

ปริญญานิพนธ์ฉบับนี้ได้รับทุนอุดหนุนการวิจัย

“กองทุน ดร.ชวาล แพรัตกุล”

มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร ประจำปี พ.ศ. 2541

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้ได้รับความอนุเคราะห์จาก รศ.ดร. บุญเชิด ภิญโญอนันตพงษ์ และ ผศ.เชาวนา ขวลิทธำรง ที่กรุณาให้คำปรึกษาแนะนำ ชี้แนะแนวทาง และแก้ไขข้อบกพร่องต่างๆ ในการทำปริญญานิพนธ์จนสำเร็จลุล่วงไปด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร ที่ได้ให้ทุนของ “กองทุน ดร.ชวาล แพร่ตกุล” แก่ผู้วิจัยในการทำปริญญานิพนธ์เล่มนี้

ขอกราบขอบพระคุณ ผู้เชี่ยวชาญทั้ง 5 ท่าน ที่ให้ความช่วยเหลือ แนะนำ แก้ไขข้อบกพร่องต่าง ๆ ในการสร้างเครื่องมือครั้งนี้

ขอขอบคุณ ผู้บริหารโรงเรียนทุกโรงเรียน นักเรียนทุกคนที่เป็นกลุ่มตัวอย่าง ในการทดลองเครื่องมือ และคณาจารย์ทุกท่าน ที่ให้ความช่วยเหลือในการเก็บรวบรวมข้อมูลครั้งนี้จนสำเร็จลุล่วงด้วยดี

ขอขอบคุณ คุณเม้ง - คุณจิตติภมล เทพครเมือง และคุณลัดดา เสือสืบพันธุ์ ที่ให้ข้อแนะนำและช่วยเหลือในการวิเคราะห์ข้อมูล

ขอขอบคุณคุณคุณประทุมพร ศรีวัฒนกุล และคุณปราณี น้อยไร่ภูมิ ที่ให้คำแนะนำและช่วยเหลือทางด้านภาษา

ขอขอบคุณ คุณอัมพร วิชัยศรี เพื่อนๆ พี่ๆ น้อง ๆ เอกการวัดผลการศึกษา และคณาจารย์โรงเรียนพระโขนงพิทยาลัย ที่เป็นกำลังใจ และให้ความช่วยเหลือในการทำปริญญานิพนธ์

ท้ายสุดนี้ ผู้วิจัยขอโน้มรำลึกถึงพระคุณของบิดา - มารดา ครู - อาจารย์ ที่ได้อบรมสั่งสอน และขอขอบคุณสมาชิกในครอบครัวทุกคนที่เป็นกำลังใจ และให้ความช่วยเหลือ สนับสนุนในทุกด้าน จนทำให้ประสบความสำเร็จในการศึกษา

อรอนงค์ บำรุง

สารบัญ

บทที่	หน้า
1 ภูมิหลัง.....	1
ความมุ่งหมายของการศึกษาค้นคว้า.....	3
ความสำคัญของการศึกษาค้นคว้า.....	4
ขอบเขตของการศึกษาค้นคว้า.....	4
นิยามศัพท์เฉพาะ.....	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	6
แบบทดสอบเลือกตอบ.....	6
ความหมายและลักษณะของการประเมินสภาพจริง.....	8
ข้อคำถามปลายเปิด.....	11
เกณฑ์การให้คะแนน.....	13
หลักฐานความเที่ยงตรงเชิงโครงสร้าง.....	15
ความเชื่อมั่นของแบบทดสอบ.....	19
งานวิจัยที่เกี่ยวข้อง.....	25
สมมติฐานของการศึกษาค้นคว้า.....	30
3 วิธีดำเนินการศึกษาค้นคว้า.....	31
ประชากร.....	31
กลุ่มตัวอย่าง.....	31
เครื่องมือที่ใช้ในการศึกษาค้นคว้า.....	31
วิธีดำเนินการสร้างเครื่องมือ.....	37
วิธีการเก็บรวบรวมข้อมูล.....	38
ขั้นตอนการวิเคราะห์เครื่องมือ.....	38
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	39
4 ผลการวิเคราะห์ข้อมูล.....	41
สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล.....	41
ผลการวิเคราะห์ข้อมูล.....	42
5 สรุป อภิปรายผล และข้อเสนอแนะ.....	53
สรุปผลการวิเคราะห์ข้อมูล.....	54
อภิปรายผล.....	55
ข้อเสนอแนะ.....	57

บทที่	หน้า
บรรณานุกรม.....	58
ภาคผนวก.....	63
ประวัติย่อผู้วิจัย	70

บัญชีตาราง

ตาราง	หน้า
1 ข้อดีและข้อจำกัดของแบบทดสอบเลือกตอบ	8
2 เมทริกซ์ หลายลักษณะ-หลายวิธี (Multitrait-Multimethod Matrix) เมื่อวัดคุณลักษณะ A และ B ด้วยวิธีวัด 1 และ 2	16
3 จำนวนห้องเรียนและจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างจำแนกตามโรงเรียน	31
4 ลักษณะเฉพาะของแบบทดสอบ	37
5 ค่าสถิติพื้นฐานของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิดและฉบับเลือกตอบ(N=50) ..	43
6 ค่าความยาก(p)และค่าอำนาจจำแนก(r)ของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด.....	44
7 ค่าความยาก(p)และค่าอำนาจจำแนก(r)ของแบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบ.....	45
8 ค่าสถิติพื้นฐาน ของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิดและฉบับเลือกตอบ(N=300)	46
9 ค่าความเที่ยงตรงเชิงโครงสร้างแบบหลายลักษณะ-หลายวิธี (Multitrait-Multimethod) ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์	47
10 ค่าความแปรปรวนร่วมระหว่างข้อคำถามของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ จำนวน 10 ข้อ.....	48
11 ส่วนเบี่ยงเบนมาตรฐานรายข้อ(S_i)น้ำหนักองค์ประกอบของคำถามแต่ละข้อในองค์ประกอบที่ k (f_{ik}) และผลคูณระหว่างส่วนเบี่ยงเบนมาตรฐานรายข้อกับน้ำหนักองค์ประกอบของคำถามแต่ละข้อในองค์ประกอบที่ k ($S_i f_{ik}$).....	49
12 ค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร ρ_{Tsk} ของไฮเซนและบอร์นสเทรท(Heise and Bohmstedt.1970).....	49
13 กำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ (λ_i^2) ผลต่างระหว่าง 1 กับกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ ($1 - \lambda_i^2$) และสัดส่วนของกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ ต่อผลต่างระหว่าง 1 กับกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ ($\lambda_i^2 / 1 - \lambda_i^2$).....	50
14 ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร Ω_w ของอัลเลน (Allen.1974).....	51
15 ค่าความแปรปรวนร่วมระหว่างคะแนนในแต่ละด้าน (λ_g^2) ผลรวมของความแปรปรวนร่วมระหว่างคะแนนในแต่ละด้าน (λ_g^2) และความแปรปรวนของคะแนนในแต่ละด้าน(S_g^2).....	51
16 ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร r_L ของเลียว (Liou.1989).....	52

บทที่ 1

บทนำ

ภูมิหลัง

การศึกษาเป็นกระบวนการที่มุ่งพัฒนาคน ให้มีความเจริญงอกงามตามทิศทางการสังคมปรารถนา โดยพยายามเปลี่ยนแปลงพฤติกรรมของผู้เรียน ให้มีคุณสมบัติตรงตามความต้องการของสังคม ดังนั้น นักการศึกษาจึงพยายามสร้างหรือปลูกฝังคุณลักษณะต่างๆ ที่จำเป็นที่ผู้เรียนยังไม่มี หรือมีอยู่น้อยให้เกิดขึ้น ในขณะที่เดียวกันก็มุ่งจัดหรือเปลี่ยนแปลงพฤติกรรมที่ไม่พึงประสงค์ให้หมดไป การจัดการศึกษาเพื่อให้บรรลุเป้าหมาย ดังกล่าว จำเป็นต้องอาศัยองค์ประกอบหลายประการ องค์ประกอบหนึ่งก็คือ การวัดผล (ไพศาล หวังพานิช. 2526 : 1-3)

เนื่องจากการวัดผลทางการศึกษาเป็นการวัดคุณสมบัติต่างๆ ของผู้เรียน เช่น ความสามารถ ความฉลาด ความถนัด เป็นต้น ซึ่งเป็นเรื่องของพฤติกรรม หรือคุณลักษณะทางจิตวิทยา ที่ต้องใช้เครื่องมือวัด และเครื่องมือวัดที่นิยมใช้กันมากในปัจจุบัน คือแบบทดสอบแบบเลือกตอบ ที่ให้นักเรียนเลือกคำตอบถูก เพียงคำตอบเดียวแล้วนำคะแนนที่ได้แต่ละข้อมารวมกัน คะแนนที่สอบได้ดังกล่าว จึงมีข้อสงสัยว่า สามารถแสดงผลการเรียนรู้ที่ซับซ้อนได้หรือไม่ คะแนนที่ได้เกิดการเรียนรู้หรือพัฒนาขึ้นจริงหรือ แบบทดสอบที่ครู ใช้เพียงพอที่จะแสดงถึงเป้าหมายสำคัญ หรือมุ่งวัดเฉพาะทักษะพื้นฐาน ทักษะที่สังเกตได้โดยตรงและทักษะ ทั่วๆ ไป จนมองข้ามความคิดขั้นสูง ความคิดขั้นซับซ้อน และวิธีการแก้ปัญหาต่างๆ ดังนั้นการใช้แบบทดสอบ เลือกตอบเพียงอย่างเดียวให้ได้ผลสอดคล้องกับการเรียนการสอนแนวใหม่ ที่มุ่งเน้นการพัฒนาคนให้มีความรู้ มีทักษะในการสื่อสาร สามารถคิดอย่างมีวิจารณญาณ แก้ปัญหาในโลกของความเป็นจริงได้อย่างมีประสิทธิภาพ และสอดคล้องกับการเรียนรู้ในโลกของความเป็นจริง จึงมีข้อจำกัดสูงมากจนเกิดความไม่มั่นใจ ต่อคะแนนสอบที่ได้ว่าจะมีความหมายแท้จริงตรงตามที่ต้องการการประเมินผล(บุญเชิด ภิญญอนันตพงษ์. 2538 : 59 ; กรมวิชาการ. 2539 : 1-3) เพื่อให้การวัดผลมีความหมายแท้จริงตามที่ต้องการวัดดังกล่าว นักการศึกษาได้เสนอแนะวิธีการประเมินผลแนวใหม่ ซึ่งลักษณะของการประเมินผลแนวใหม่มีดังนี้ 1) ด้าน พฤติกรรม เน้นให้แสดงวิธีคิด ลงมือทำมากกว่าการเลือกตอบและใช้ทักษะบูรณาการมากกว่าการแยกแยะ ทักษะ 2) เน้นการเขียนตอบตามสภาพจริงในสิ่งที่มีความหมายต่อนักเรียน และเน้นทักษะที่ซับซ้อนเพื่อวัด ความงอกงามและความก้าวหน้า 3) เน้นการประเมินหลายมิติทั้งความสามารถรวม และความสามารถพิเศษ ค้นหาคำความสามารถเด่นของนักเรียน เปิดโอกาสให้นักเรียนได้พัฒนาและแสดงความสามารถพิเศษนั้น 4) เน้นการประเมินเป็นกลุ่มโดยใช้ทักษะกระบวนการกลุ่ม หรือผลงานที่ทำร่วมกันเป็นกลุ่ม (บุญเชิด ภิญญอนันตพงษ์. 2538) หรือมีลักษณะดังนี้ 1) เน้นกระบวนการและผลของการกระทำ 2) เน้นความสามารถจริงในการดำเนินชีวิต 3) สะท้อนให้เห็นถึงการใช้ความคิดพิจารณาไตร่ตรอง การใช้เหตุผล และการแก้ปัญหา 4) จำลองสภาพชีวิตจริงในชุมชน และ 5) เน้นความพึงพอใจในสภาพของตน (ส.วาสนา ประวาลพฤษย์. 2539 : 2)

จากลักษณะของการประเมินผลดังกล่าว นักการศึกษาจึงได้พยายามค้นหารูปแบบการประเมินมา เสริมเพิ่มเติม เพื่อให้การประเมินผลมีความสมบูรณ์ยิ่งขึ้น วิธีที่น่าสนใจวิธีหนึ่งคือ การประเมินผลจากสภาพ จริงที่เรียกว่า Authentic Assessment (AA) หรือ Authentic Performance Assessment (APA) ซึ่ง เป็นการวัดและประเมินผล ที่บอกถึงระดับความรู้ความสามารถที่เกิดขึ้นในตัวผู้เรียน อันเนื่องมาจากการให้

นักเรียนได้ลงมือปฏิบัติ จนกระทั่งได้ชิ้นงานในรูปแบบใดรูปแบบหนึ่งออกมาในตอนสุดท้าย เทคนิคที่นิยมใช้ตามแนวการประเมินผลจากสภาพจริงอย่างหนึ่ง คือการใช้แบบทดสอบวัดความสามารถจริง (Authentic Performance Test) ซึ่งลักษณะของข้อสอบจะเป็นข้อคำถามที่เป็นปัญหาให้ผู้เรียนคิดและเขียนตอบเอง โดยปัญหาจะต้องมีความหมายต่อผู้เรียน มีความสำคัญเพียงพอ และเลียนแบบสภาพจริงในชีวิตของผู้เรียนเหมาะสมกับวัยและสภาพชีวิตของเขา ในการสอบผู้สอบจะต้องใช้ความรู้ความสามารถหลายด้าน ต้องคิดอย่างลึกซึ้ง มีวิธีการและขั้นตอนที่ชัดเจนและเหมาะสม คำตอบถูกมีหลายคำตอบ วิธีการหาคำตอบมีหลายวิธี (ส.วาสนา ประมวลพฤกษ์. 2539 : 5) ข้อคำถามปลายเปิด (open-ended question) นับว่าเป็นเครื่องมือชนิดหนึ่งที่ถูกนักถึงจากนักปฏิรูปการศึกษาว่า สามารถใช้วัดกระบวนการคิดขั้นสูงและการให้เหตุผลที่สะท้อนให้เห็นถึง การใช้สติปัญญา และความสามารถในการปฏิบัติของผู้สอบได้ดีกว่าแบบทดสอบเลือกตอบ ถ้าข้อคำถามนั้นเป็นตัวแทนที่ดีของเนื้อหาและท้าทายให้นักเรียนแสดงความคิดที่คาดหวังหรือเป็นไปได้ของงานสอบนั้น ๆ (Linn, Baker, and Dunbar. 1991 : 15-21)

คณิตศาสตร์เป็นวิชาที่เกี่ยวข้องกับความคิด กระบวนการ และเหตุผล แต่การจัดการเรียนการสอนครูมักสอนให้เด็ก "จำ" และ "ทำ" ตามคำบอกของครูมากกว่าการประยุกต์ใช้ หรือด้วยการสอนแบบอธิบายและแสดงเหตุผล คือขณะที่ผู้สอนอธิบายก็จะวิเคราะห์ ชี้แจง ที่ความ ให้ผู้เรียนเข้าใจ และสรุปด้วยตนเอง ผู้เรียนเป็นผู้ฟังเสียส่วนใหญ่และไม่มีโอกาสร่วมกิจกรรมการเรียนการสอนมากนัก (ยุพิน พิพิธกุล. 2527 : 122-123) นอกจากนี้ยังนิยมวัดผลประเมินผลเพื่อดูคุณภาพโดยรวม เพื่อการตัดสินใจในระดับต่างๆ หรือจากเกณฑ์ผ่านและไม่ผ่านด้วยการสอบอย่างมีระบบเป็นขั้นตอน เน้นการสอบด้วยแบบทดสอบปรนัยเลือกตอบ ซึ่งไม่ให้คุณค่าในการพัฒนาการเรียนการสอนเท่าที่ควร ซึ่งซิลเวอร์ (Magone, et al. 1994 : 317 citing Silver. 1992) กล่าวว่า การใช้แบบทดสอบเลือกตอบในวิชาคณิตศาสตร์ไม่ก่อให้เกิดทักษะด้านการคิด มักเน้นความรู้ความจำมากกว่าความรู้ที่แท้จริงและการประยุกต์ใช้ในชีวิตประจำวัน นอกจากนี้สมาคมครูคณิตศาสตร์ในสหรัฐอเมริกา (The National Council of Teachers of Mathematics : NCTM .1989 : 23) กล่าวไว้ว่าการใช้แบบทดสอบเลือกตอบในวิชาคณิตศาสตร์ ขัดกับจุดมุ่งหมายการเรียนการสอนวิชาคณิตศาสตร์ที่เน้นความสำคัญด้านการคิด ความเข้าใจ การให้เหตุผล การแก้ปัญหาที่เน้นความต่อเนื่อง การประยุกต์ใช้ และการสื่อสาร โดยให้ความสามารถในการแก้ปัญหาเป็นเป้าหมายและจุดเน้นที่สำคัญของการเรียนรู้คณิตศาสตร์ และให้การแก้ปัญหาเป็นส่วนสำคัญในกิจกรรมการเรียนการสอน สำหรับประเทศไทยแม้ว่าจุดประสงค์ของหลักสูตรคณิตศาสตร์ ระดับมัธยมศึกษาตอนต้นพุทธศักราช 2521 (ฉบับปรับปรุง พุทธศักราช 2533) ได้กำหนดไว้อย่างชัดเจนว่า 1) เพื่อให้นักเรียนมีความรู้ความเข้าใจในวิชาคณิตศาสตร์ 2) สามารถคิดอย่างมีเหตุผล ใช้เหตุผลในการแสดงความคิดเห็นอย่างมีระบบชัดเจนและรัดกุม 3) ให้มีทักษะในการคิดคำนวณ เห็นประโยชน์ของวิชาคณิตศาสตร์ทั้งที่มีต่อชีวิตประจำวัน และเป็นเครื่องมือในการแสวงหาความรู้ และ 4) สามารถนำความรู้ ความเข้าใจ และทักษะทางคณิตศาสตร์ไปใช้ในชีวิตประจำวัน และเป็นพื้นฐานในการศึกษา (กรมวิชาการ. 2535 : 40) แต่การเรียนการสอนที่ผ่านมายังไม่บรรลุผลเท่าที่ควร ดังจะเห็นได้จากผลการวิจัยเรื่อง ประสิทธิภาพของการมัธยมศึกษา คณะแผนกสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ของกลุ่มตัวอย่างที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ได้คะแนนร้อยละ 31.53 และนักเรียนชั้นมัธยมศึกษาปีที่ 6 ได้คะแนนร้อยละ 29.63 (กรมวิชาการ. 2538 : 28,88) ซึ่งสอดคล้องกับผลงานวิจัยของไทยร่วมกับนานาชาติเกี่ยวกับวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 พบว่าผลสัมฤทธิ์ในวิชาคณิตศาสตร์ของนักเรียนไทยยังอยู่ในระดับที่ควรปรับปรุงให้ดีขึ้น (ภัทรกุล จิรวิทย์. 2530 : 11) และจากผลการวิจัยของกองวิจัยทางการศึกษากกรมวิชาการ (2532) พบว่า ลักษณะข้อบกพร่องของเด็กทางคณิตศาสตร์มี 3

ลักษณะ คือ ด้านความคิดรวบยอด ด้านการคิดคำนวณ และด้านการแก้ปัญหาโจทย์ ทั้งนี้อาจเนื่องมาจากเป้าหมายของการศึกษามักกำหนดความสามารถที่ติดอยู่กับรายวิชา ไม่ได้เลยไปถึงการนำความรู้ไปใช้ จึงมุ่งสอนให้จบเนื้อหา ขาดการเชื่อมโยงในการนำความรู้ไปประยุกต์ใช้ ผู้เรียนไม่ค่อยมีโอกาสร่วมคิด ร่วมทำ ร่วมแก้ปัญหา (โกวิท ประวาลพุกษ์, 2534 : 3) เพื่อให้การวัดผลประเมินผลวิชาคณิตศาสตร์ สอดคล้องกับจุดมุ่งหมายดังกล่าว ได้มีหน่วยงานต่างๆทั้งในประเทศและนอกประเทศพยายามค้นหารูปแบบการประเมินผลเพื่อลดบทบาทการใช้ข้อสอบมาตรฐานแบบเลือกตอบ QUASAR (Quantitative Understanding : Amplifying Student Achievement and Reasoning) เป็นโครงการหนึ่งของประเทศสหรัฐอเมริกา ที่จัดทำขึ้นเพื่อปรับปรุงการเรียนการสอนวิชาคณิตศาสตร์ ของนักเรียนระดับมัธยมศึกษาตอนต้นในประเทศสหรัฐอเมริกา โดยเน้นที่เครื่องมือประเมินผลทางด้านพุทธิพิสัย (Cognitive Assessment Instrument) ที่เรียกว่า QCAI (The Quasar Cognitive Assessment Instrument) ซึ่งประกอบด้วย ชุดของงานสอบปลายเปิด (Open-ended Tasks) ที่มีลักษณะดังนี้ 1)นักเรียนเลือกหรือหาคำตอบ โดยมีคำอธิบายหรือเหตุผลกำกับว่าทำไมถึงเลือกหรือหาคำตอบเช่นนั้น 2) ข้อคำถามต้องมีความซับซ้อนสามารถหาคำตอบได้หลายคำตอบ ในรูปของการเขียนบรรยาย ใช้รูปภาพ กราฟ ตาราง การแจกแจงกรณีที่เป็นไปได้ หรือสัญลักษณ์ทางคณิตศาสตร์ และ3)ความซับซ้อนนั้นจะเหมาะสม สำหรับวัดความก้าวหน้า ของนักเรียนในช่วงเวลาหนึ่ง (Magone. et al. 1994 : 319) จากเหตุผลและแนวคิดข้างต้น แสดงถึงแนวโน้มของการกลับมาใช้แบบทดสอบปลายเปิด ซึ่งมีลักษณะคล้ายคลึงกับแบบทดสอบอัตนัย ผู้วิจัยสนใจที่จะศึกษาว่าแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ มีความเที่ยงตรงเชิงโครงสร้างเป็นอย่างไร โดยพิจารณาจากหลักฐานความเที่ยงตรงที่คำนวณหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) และที่คำนวณด้วยสูตรของ ไฮเซน และบอร์นสเตดท์ (Heise and Bohmstedt. 1970 : 104-118) และพิจารณาค่าความเชื่อมั่นจากสัมประสิทธิ์ Ω_w ของอัลเลน (Allen. 1974) และสัมประสิทธิ์ความสอดคล้องภายใน ซึ่งคำนวณจากสูตรความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ ตามสูตรของเลียว (Liou) โดยใช้เนื้อหาวิชาคณิตศาสตร์(ค 203) ชั้นมัธยมศึกษาปีที่ 2 ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) เรื่อง ระบบจำนวนเต็ม เศษส่วน และทศนิยม

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ สองวิธี ดังนี้

1.1 วิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait-Multimethod)

1.2 วิธีหาความสัมพันธ์ระหว่างคะแนนจริงกับคะแนนองค์ประกอบ ที่คำนวณด้วยสูตร

$P_{r_{sk}}$ ของไฮเซน และ บอร์นสเตดท์ (Heise and Bohmstedt. 1970)

2. เพื่อหาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ สองวิธี ดังนี้

2.1 วิธีวิเคราะห์น้ำหนักองค์ประกอบ ที่คำนวณด้วยสูตร Ω_w ของอัลเลน (Allen. 1974)

2.2 วิธีวิเคราะห์แบบคะแนนจริงสัมพันธ์ ที่คำนวณด้วยสูตร r_{12} ของเลียว (Liou. 1989)

ความสำคัญของการศึกษาค้นคว้า

การศึกษาค้นคว้าครั้งนี้ เป็นการศึกษาคุณภาพด้านความเที่ยงตรงและความเชื่อมั่น ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ พร้อมเกณฑ์การให้คะแนน ผลที่ได้จากการศึกษาจะเป็นแนวทางในการสร้างแบบทดสอบปลายเปิดที่มีคุณภาพ และนำไปใช้ในการวัดผลการเรียนวิชาคณิตศาสตร์ต่อไป

ขอบเขตของการศึกษาค้นคว้า

1. ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียน ชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 ของโรงเรียนในสังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพฯ ที่เป็นโรงเรียนสหศึกษา จำนวน 11 โรงเรียน 159 ห้องเรียน จำนวนนักเรียน 6,987 คน
2. กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 สังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพฯ ที่เป็นโรงเรียนสหศึกษา จำนวน 300 คน
3. เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นแบบทดสอบวิชาคณิตศาสตร์ 2 ฉบับ คือ
 - 3.1 แบบทดสอบปลายเปิด
 - 3.2 แบบทดสอบเลือกตอบ

นิยามศัพท์เฉพาะ

1. แบบทดสอบปลายเปิด (Open - ended Questions) หมายถึง แบบทดสอบที่ประกอบด้วยข้อคำถาม ที่มีลักษณะเป็นโจทย์ที่เน้นความคิดซับซ้อน เลียนแบบสภาพความเป็นจริงในชีวิตประจำวัน ทำให้นักเรียนแก้ปัญหาโดยใช้ความรู้ความสามารถหลายด้านที่ต้องคิดอย่างลึกซึ้ง มีคำตอบที่เป็นไปได้หลายคำตอบ มีวิธีการหาคำตอบที่หลากหลาย
2. แบบทดสอบปลายเปิดวิชาคณิตศาสตร์ หมายถึงแบบทดสอบที่ประกอบด้วยชุดของข้อคำถามปลายเปิดที่ให้นักเรียนอธิบาย หรือแสดงเหตุผลในการตอบ หรือแสดงวิธีหาคำตอบ ที่มีวิธีการหรือขั้นตอนที่ชัดเจน ให้คะแนนตามเกณฑ์รูบริค (Rubric) ที่กำหนดไว้ และตอบในกระดาษคำตอบที่กำหนดให้ ซึ่งวัดด้านต่าง ๆ ดังนี้
 - 1.1 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์
 - 1.2 การแก้ปัญหา
 - 1.3 การให้เหตุผล
3. แบบทดสอบเลือกตอบ (Multiple Choice) หมายถึง แบบทดสอบที่ข้อคำถามแต่ละข้อสร้างมาจากข้อคำถามของแบบทดสอบปลายเปิด ข้อคำถามแต่ละข้อ จะมี 5 ตัวเลือก เป็นตัวเลือกถูกเพียงคำตอบเดียว ที่วัดด้านต่าง ๆ ดังนี้
 - 3.1 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์
 - 3.2 การแก้ปัญหา
 - 3.3 การให้เหตุผล
4. ความเข้าใจในความคิดรวบยอด หมายถึง ความสามารถในการบ่งบอก อธิบาย หลักการ สูตร กฎเกณฑ์ทางคณิตศาสตร์และการนำไปใช้

5. การแก้ปัญหา หมายถึง ความสามารถในการคิดหาคำตอบโดยอาศัย ความคิดรวบยอด กฎ สูตร หลักการ ทักษะ และการดำเนินการทางคณิตศาสตร์

6. การให้เหตุผล หมายถึง การแสดงแนวคิดเกี่ยวกับการสร้างหลักเกณฑ์ การหาความสัมพันธ์ และการสรุปผลจากข้อเท็จจริงที่กำหนดให้

7. หลักฐานความเที่ยงตรงของแบบทดสอบ (Validity) หมายถึง ข้อมูลที่สนับสนุนหรือแสดงว่าแบบทดสอบสามารถวัดในสิ่งที่ต้องการวัดได้ตรงตามจุดมุ่งหมาย ในการวิจัยในครั้งนี้ แสดงหลักฐานความเที่ยงตรงของแบบทดสอบ โดยวิธีต่อไปนี้

7.1 หลักฐานความเที่ยงตรงตามเนื้อหา หมายถึง ข้อคำถามของแบบทดสอบที่แสดงว่าเป็นตัวแทนที่แท้จริงของเนื้อหาตามที่กำหนดไว้ในตารางวิเคราะห์หลักสูตรภาคเนื้อหา ในการวิจัยครั้งนี้ พิจารณาความเกณฑ์ ความคิดซับซ้อน (Cognitive Complexity) และขอบเขตเนื้อหา (Content quality) ด้วยการวิเคราะห์ทางตรรกศาสตร์ ของเนื้อหางานสอบ และพฤติกรรมที่คาดหวัง ตรวจสอบโดยใช้ผู้เชี่ยวชาญ

7.2 หลักฐานความเที่ยงตรงเชิงโครงสร้าง หมายถึง ข้อคำถามของแบบทดสอบที่แสดงว่าสามารถวัดได้ตรงตามลักษณะโครงสร้าง ในการวิจัยครั้งนี้ คำนวณหาความสัมพันธ์หลายลักษณะ-หลายวิธี และหาความสัมพันธ์ระหว่างคะแนนจริงกับคะแนนองค์ประกอบ ดังนี้

7.2.1 หาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) หมายถึง การแสดงความเที่ยงตรงเชิงโครงสร้างที่มีลักษณะที่วัดหลายลักษณะ และวิธีวัดหลายวิธี แบ่งความเที่ยงตรงออกเป็น 2 ชนิด ดังนี้

7.2.1.1 ความเที่ยงตรงเชิงเหมือน (Convergent Validity) หมายถึง การวัดในลักษณะเดียวกันด้วยแบบทดสอบต่างชนิดกัน จะมีค่าสหสัมพันธ์กันสูง ในการวิจัยครั้งนี้ค่าสหสัมพันธ์จะสูงกว่าความเที่ยงตรงเชิงจำแนก (Allen and Yen. 1979 : 109-110 ; citing Cambell and Fiske. 1959)

7.2.1.2 ความเที่ยงตรงเชิงจำแนก (Discriminant Validity) หมายถึง การวัดในลักษณะต่างกันด้วยแบบทดสอบชนิดเดียวกันหรือต่างชนิดกัน จะมีค่าสหสัมพันธ์กันต่ำ ในการวิจัยครั้งนี้ค่าสหสัมพันธ์จะต่ำกว่าความเที่ยงตรงเชิงเหมือน (Allen and Yen. 1979 : 109-110 ; citing Cambell and Fiske. 1959)

7.2.2 หาความสัมพันธ์ระหว่างคะแนนจริงในแต่ละคุณลักษณะย่อยของแบบทดสอบ กับคะแนนองค์ประกอบ (Composite Score) ทั้งหมดของแบบทดสอบ ที่คำนวณด้วยสูตรของ ไฮเซน และ บอร์นสเตดท์ (Haise and Bohmstedt)

8. ความเชื่อมั่นของแบบทดสอบ (Reliability) หมายถึง สัดส่วนของความแปรปรวนของคะแนนจริง ต่อความแปรปรวนของคะแนนที่สอบได้ ในการวิจัยครั้งนี้ หาค่าความเชื่อมั่นของแบบทดสอบจาก

8.1 วิธีวิเคราะห์น้ำหนักองค์ประกอบ ที่คำนวณด้วยสูตร Ω_w ของอัลเลน (Allen.1974)

8.2 สัมประสิทธิ์ความสอดคล้องภายใน ซึ่งคำนวณจากความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ตามสูตร r_{12} ของเลียว (Liu.1988)

9. การตรวจให้คะแนน ใช้แบบประเมินรวม (Holistic Score) จากคะแนนการตอบของนักเรียนแต่ละข้อ เมื่อใช้เกณฑ์การให้คะแนนแบบ รูบริค (Rubric)

10. เกณฑ์การให้คะแนน หมายถึง ตัวเลขที่กำหนดขึ้นเพื่อบอกระดับความสามารถในการสอบแต่ละข้อ

11. ผู้เชี่ยวชาญ หมายถึง ผู้ที่มีวุฒิทางการศึกษาตั้งแต่ระดับปริญญาโทขึ้นไป ทางด้านการวัดผลการศึกษาหรือทางด้านคณิตศาสตร์ หรือเป็นผู้ที่มีประสบการณ์การสอนวิชาคณิตศาสตร์ระดับมัธยมศึกษา ไม่น้อยกว่า 5 ปี จำนวน 5 ท่าน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้ทำการศึกษานโยบายต่าง ๆ ตามลำดับ ต่อไปนี้

1. แบบทดสอบเลือกตอบ
2. ความหมายและลักษณะของการประเมินสภาพจริง
3. ข้อคำถามปลายเปิด
4. เกณฑ์การให้คะแนน
5. ความเที่ยงตรงเชิงโครงสร้าง
6. ความเชื่อมั่นของแบบทดสอบ
7. งานวิจัยที่เกี่ยวข้อง

1. แบบทดสอบเลือกตอบ

แบบทดสอบเลือกตอบ เข้ามามีบทบาทในระบบการศึกษาของไทยมากกว่า 80 ปีแล้ว โดยมีวัตถุประสงค์ ที่จะให้การสอบต่าง ๆ สามารถวัดได้ครอบคลุม ยุติธรรม และให้ความเชื่อมั่นมากขึ้นกว่าเดิม (ชวาล แพทย์กุล, 2530 : 52)

ลักษณะของแบบทดสอบเลือกตอบ

ข้อสอบเลือกตอบประกอบด้วยส่วนสำคัญ 2 ส่วน คือ

1. ส่วนที่เป็นปัญหา หรือคำถาม(Stem) คือข้อความที่เขียนเป็นประโยคคำถามที่สมบูรณ์หรือไม่สมบูรณ์ ทำหน้าที่เป็นสิ่งที่เร้า เพื่อให้ผู้สอบตอบสนอง
2. ส่วนที่เป็นตัวเลือก (Alternative or choice or option) ซึ่งจะมีทั้งตัวเลือกที่เป็นตัวถูก(key) และตัวเลือกที่เป็นตัวลวง (foils หรือ distractors) (อนันต์ ศรีโสภณ .2524 : 177)

การสร้างแบบทดสอบเลือกตอบ

ชวาล แพทย์กุล(2516 : 116-190) อนันต์ ศรีโสภณ (2524 : 126-134) วิทยุญา วิชาลาภรณ์ (2530 : 152-165) ธอนไดค์ (Thomdike. 1971 : 113-119) และ ไอเคน (Aiken. 1987 : 44-58) ได้เสนอหลักการสร้างแบบทดสอบเลือกตอบไว้คล้ายกัน ดังนี้

หลักการสร้างข้อคำถาม

1. เขียนคำถามให้เป็นประโยคคำถามที่สมบูรณ์ ไม่เขียนแบบลักษณะต่อความ
2. เน้นเรื่องที่จะถามให้ชัดเจน ตรงจุด อ่านแล้วเข้าใจทันทีว่าถามอะไร
3. ถามข้อละปัญหา แต่ละข้อเป็นอิสระจากกัน
4. ใช้ภาษาให้เหมาะสมกับระดับของผู้สอบ ไม่ใช่คำฟุ่มเฟือย หลีกเลี่ยงคำว่า เสมอ ไม่เคย และ

ทั้งหมด

5. ไม่ควรใช้คำถามปฏิเสธ หรือปฏิเสธซ้อน ถ้าจำเป็นต้องใช้ควรขีดเส้นใต้คำว่า "ไม่"
6. ควรถามให้ใช้ความคิด ยั่วให้เกิดความคิด หลีกเลี่ยงคำถามที่คล่องปากของนักเรียน

หลักการสร้างตัวเลือก

1. ให้คำตอบถูกมีเพียงคำตอบเดียว
2. ควรจัดเรียงตัวเลือกให้เป็นระเบียบ
3. ตัวเลือกทุกตัวควรจะเป็นอิสระแก่กัน
4. ตัวเลือกควรจะมีควมยาวใกล้เคียงกัน
5. ควรกระจายตำแหน่งของคำตอบถูก
6. เลี่ยงการใช้ตัวเลือก "ถูกหมดทุกข้อ"
7. ตัวเลือกอยู่ในกลุ่มเดียวกัน หรือมีความเป็นเอกพันธ์
8. หลีกเลี่ยงคำที่ชี้แนะคำตอบ ใช้ตัวชี้ในตัวเลือก
9. ควรใช้ตัวเลือกปลายเปิดให้เหมาะสมกับเนื้อหาวิชาและสถานการณ์
10. ตัวเลือกทุกตัวควรมีทำที่ว่าจะถูก ไม่ควรใช้ตัวเลือกที่ผิดเด่นชัดมากเกินไป
11. จำนวนตัวเลือกควรมี 4-5 ตัวเลือก หรือพิจารณาให้เหมาะสมกับวัย ระดับชั้นของนักเรียน
12. ตัวเลือกควรจะเป็นคำที่สามารถต่อความกับตัวคำถามได้อย่างสอดคล้องและถูกหลักภาษา

ใช้ภาษาที่เหมาะสมกับผู้เรียน หลีกเลี่ยงการใช้คำที่คัดลอกจากตำรา

ชาวล แพริทกุล (2530 : 60) กล่าวว่า การสร้างตัวเลือกของแบบทดสอบเลือกตอบ นิยมใช้ตัวเลือกที่ใกล้เคียงกัน และเป็นเอกพันธ์ (Homogeneous) เพื่อให้เด็กคิดหาความสัมพันธ์ หรือนำมาเปรียบเทียบกัน และมีโอกาสจะถูกเลือกเท่า ๆ กันทั้งนี้เพื่อพรางตาไม่ให้นักเรียนเดา และให้คำถามนั้นเป็นสิ่งจูงใจให้นักเรียนใช้ความคิดวินิจฉัยชี้ขาดว่า ตัวเลือกใดเป็นตัวเลือกที่เหมาะสมที่สุดจากตัวเลือกทั้งหมดที่ให้ไว้ และ นอล (Noll. 1957 : 130) กล่าวว่า ตัวลวงที่ดีต้องเป็นตัวลวงที่ทำหน้าที่นี้ได้จริง และมีความใกล้เคียงกับตัวถูกมากที่สุด เพื่อให้นักเรียนที่ไม่มีความรู้จริงในเรื่องนั้น เลือกตอบตัวลวงเหล่านั้น ซึ่งสอดคล้องกับความคิดเห็นของโพเวล (Powell. 1968 : 403-413) ที่กล่าวว่าแบบทดสอบที่ดี ตัวเลือกทุกตัวควรมีโอกาสถูกเลือกพอ ๆ กัน

การตรวจให้คะแนน

ข้อสอบแบบเลือกตอบสามารถตรวจได้ง่าย รวดเร็ว และมีความเป็นปรนัย เพราะใช้คีย์เฉลย(key)

ตรวจ

ข้อดีและข้อจำกัดของแบบทดสอบเลือกตอบ

บุญเชิด ภิบุญอนันตพงษ์ ได้กล่าวถึงข้อดี และข้อจำกัดของแบบทดสอบเลือกตอบไว้ดังนี้

(บุญเชิด ภิบุญอนันตพงษ์. 2525 : 72-73)

ตาราง 1 ข้อดีและข้อจำกัดของแบบทดสอบเลือกตอบ

ข้อดี	ข้อจำกัด
1. สามารถวัดผลผลิตทางการเรียนรู้ได้หลายด้าน ตั้งแต่ความรู้ความจำ จนถึงกระบวนการทางปัญญาขั้นสูง	1. สร้างลำบากมากเพราะต้องคิดตัวเลือกให้เหมาะสม และใช้เวลาสร้างนาน
2. มีความแม่นยำสูง เพราะสามารถวัดสิ่งที่ต้องการวัดได้ครอบคลุม	2. ใช้เวลาในการตอบนานกว่าข้อสอบประเภทอื่น ๆ
3. มีความเชื่อมั่นสูงกว่าแบบทดสอบอื่น ๆ เพราะมีความเป็นปรนัยและมีผลจากการเดาถูกน้อย	3. นักเรียนเก่งมักมีแนวโน้มทำคะแนนได้มาก
4. สามารถควบคุมระดับความยากง่ายของข้อสอบได้โดยการเขียนตัวลงไปเป็นเอกพจน์	4. ไม่เหมาะที่จะวัดความคิดริเริ่ม หรือความคิดสร้างสรรค์
5. ตรวจให้คะแนนง่ายและรวดเร็ว เป็นปรนัยสามารถใช้เครื่องตรวจได้	5. ไม่สามารถวัด การรวบรวมความคิด การเสนอความคิด และทักษะการเขียน
6. สามารถวินิจฉัยข้อบกพร่อง หรือความไม่เข้าใจเนื้อหา ได้โดยพิจารณาจากตัวลง	6. สิ้นเปลืองค่าใช้จ่ายสูง

2. ความหมายและลักษณะของการประเมินสภาพจริง

การวัดผลประเมินผล เป็นองค์ประกอบที่สำคัญยิ่งในโปรแกรมการเรียนการสอน ที่ใช้ในการตรวจสอบผลการเรียนว่า ผู้เรียนเกิดคุณลักษณะต่าง ๆ ตามที่กำหนดไว้มากน้อยเพียงใด มีสิ่งใดบกพร่องควรแก่การแก้ไข (ไพศาล หวังพานิช. 2526 : 3) ซึ่งเดิมการวัดผลมักจะแยกออกจากการเรียนการสอน เน้นการสอบปลายภาคและการสอบเลื่อนชั้นด้วยข้อสอบเลือกตอบเป็นส่วนใหญ่ ซึ่งทำให้นักเรียนขาดโอกาสแสวงหาความรู้ด้วยตนเอง ขาดการใช้กระบวนการคิดวิเคราะห์วิจารณ์ รวมทั้งขาดความสามารถในการใช้ภาษาเพื่อการสื่อความหมายให้สมบูรณ์ ทำให้ไม่สามารถประเมินความสามารถที่แท้จริงได้อย่างครบถ้วน ก่อให้เกิดการพัฒนาองค์กรแห่งการเรียนรู้ เพื่อลดบทบาทการใช้แบบทดสอบเลือกตอบและแสวงหาทางเลือกในการประเมินผลที่สามารถประเมินความสามารถที่แท้จริงของผู้เรียน การประเมินสภาพจริง (Authentic Assessment) นับเป็นทางเลือกหนึ่งที่ถูกนักถึงจากนักปฏิรูปการศึกษา (กรมวิชาการ. 2539 : 3-5)

ความหมายของการประเมินสภาพจริง

กรมวิชาการ (2539 : 11) กล่าวว่า การประเมินสภาพจริง หมายถึง กระบวนการสังเกต บันทึก และรวบรวมข้อมูลจากงานและวิธีการที่นักเรียนทำ เพื่อเป็นพื้นฐานของการตัดสินใจในการศึกษาถึงผลกระทบต่อเด็กเหล่านั้น เน้นการประเมินทักษะการคิดที่ซับซ้อนในการทำงานของนักเรียน ความสามารถในการแก้ปัญหา และการแสดงออกที่เกิดจากการปฏิบัติในสภาพจริงในการเรียนการสอน ที่เน้นนักเรียนเป็นศูนย์กลาง

สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ (2539 : 4) กล่าวว่า การประเมินสภาพจริงเป็นการประเมินการกระทำ การแสดงออกหลาย ๆ ด้านของนักเรียนตามสภาพความเป็นจริงทั้งในและนอกห้องเรียน มีลักษณะเป็นการประเมินอย่างไม่เป็นทางการ

กลาตเลอร์ (เจริญ คำยัง. ม.ป.ป. : 2 ; อ้างอิงจาก Gradler.1991) กล่าวว่า การประเมินสภาพจริง เป็นการประเมินที่ไม่เน้นการประเมินทักษะเบื้องต้น (Skill Assessment) แต่จะเน้นการประเมินทักษะการคิดอย่างซับซ้อนของนักเรียน (Complex Thinking Skill) ความร่วมมือในการแก้ปัญหา และการประเมินตนเอง ซึ่งคล้ายกับโลกของงานจริง(World of Work) ภายนอกโรงเรียน

จัสมีน (เจริญ คำยัง. ม.ป.ป. : 2 ; อ้างอิงจาก Jasmine.1992) กล่าวว่า การประเมินสภาพจริง เป็นการสะท้อนให้เห็นการสังเกตสภาพงานปัจจุบันของนักเรียน และสิ่งที่นักเรียนได้ปฏิบัติจริง

นิวแมน และ วีกินส์ (เจริญ คำยัง. ม.ป.ป. : 3 ; อ้างอิงจาก Newman.1988 และ Wiggins.1989) กล่าวว่า การประเมินสภาพจริงเป็นลักษณะของการให้นักเรียนแก้ปัญหา การเขียนเรียงความ การกระทำเพื่อแก้ปัญหาทางวิทยาศาสตร์ การสร้างสถานการณ์ให้เหมือนการแก้ปัญหาในโลกแห่งความเป็นจริง และการรวบรวมงานของนักเรียนในแหล่งรวมข้อมูล

นิตโค (Nitko.1996 : 255 ; citing Baron.1991,Howath.1991 and Jones.1994) กล่าวว่า การประเมินสภาพจริงประกอบด้วยลักษณะเด่น 4 ประการ คือ

1. เน้นการประยุกต์ใช้ หมายถึง การประเมินว่าอะไรคือสิ่งที่ผู้เรียนสามารถทำได้จากเงื่อนไขที่ประเมิน
2. เน้นการประเมินทางตรง หมายถึง การประเมินจุดมุ่งหมายการเรียนรู้ทางตรงเปรียบเทียบกับ การเรียนรู้ทางอ้อม
3. ใช้ปัญหาจากสภาพจริง หมายถึง การกำหนดขอบเขตของงานตามสภาพความเป็นจริงเพื่อที่นักเรียนจะสามารถนำความรู้ไปใช้ในชีวิตประจำวันได้
4. ทำทนายให้นักเรียนแสดงความคิดอย่างอิสระหมายถึง ขอบเขตของงานจะต้องทำทนายให้นักเรียนแสดงความคิด มากกว่าการหาคำตอบที่ถูกต้อง เน้นการทำงานเป็นกลุ่มเพื่อให้ได้คำตอบที่สมบูรณ์

หลักการประเมินผลจากสภาพจริง

กรมวิชาการ (2539 : 27) ได้เสนอแนะหลักการประเมินสภาพจริง ไว้ดังนี้

1. เป็นการประเมินความก้าวหน้าและการแสดงออกของนักเรียนแต่ละคน (มิใช่เปรียบเทียบกับกลุ่ม) บนรากฐานของทฤษฎีทางพฤติกรรมศาสตร์ และด้วยเครื่องมือประเมินที่หลากหลาย
2. การประเมินผลจากสภาพจริง จะต้องมีการฐานบนพัฒนาการ และการเรียนรู้ทางสติปัญญาที่หลากหลาย
3. การประเมินผลจากสภาพจริงและการพัฒนาหลักสูตรที่เหมาะสม จะต้องจัดทำให้ส่งเสริมซึ่งกันและกัน คือ จะต้องพัฒนามาจากบริบทที่มีรากฐานทางวัฒนธรรมที่นักเรียนอาศัยอยู่ และต้องเรียนรู้ให้ทันกับกระแสการเปลี่ยนแปลงของโลก
4. ความรู้ในเนื้อหาสาระทั้งในทางกว้างและลึก จะนำไปสู่การพัฒนาให้นักเรียนเรียนรู้มากขึ้น เพื่อให้ผู้เรียนบรรลุเป้าหมาย และเสริมสร้างศักยภาพของผู้เรียนอย่างเต็มที่
5. การเรียน การสอน การประเมิน จะต้องหลอมรวมกันและการประเมินต้องประเมินต่อเนื่องตลอดเวลาที่ทำการเรียนการสอน โดยผู้เรียนมีส่วนร่วม
6. การเรียน การสอน การประเมิน เน้นการปฏิบัติจริงในสภาพที่สอดคล้องหรือใกล้เคียงกับธรรมชาติความเป็นจริงของการดำเนินชีวิต เปิดโอกาสให้ผู้เรียนได้คิดงานด้วยตนเอง
7. การเรียน การสอน จะต้องเป็นไปเพื่อการพัฒนาศักยภาพอย่างเต็มที่สูงสุดตามสภาพที่เป็นจริงของแต่ละบุคคล

ลักษณะสำคัญของการประเมินสภาพจริง

ธนารักษ์ ปิ่นเทียน (2540 : 5) กล่าวถึง ลักษณะสำคัญของการประเมินสภาพจริง ไว้ดังนี้

1. เป็นการประเมินผลตลอดเวลา (Kid Watching) กระทำได้ตลอดเวลาทุกสถานการณ์ ทั้งที่บ้าน โรงเรียน และชุมชน สังเกตพฤติกรรมต่าง ๆ โดยใช้การตัดสินใจของมนุษย์ในการให้คะแนน ทั้งนี้เพื่อพัฒนาการด้านต่าง ๆ ของนักเรียน
2. อยู่บนพื้นฐานของการปฏิบัติจริง เน้นสภาพปัญหาที่สอดคล้องกับความเป็นจริงในชีวิตประจำวัน เน้นการแก้ปัญหาที่สะท้อนชีวิตจริง
3. กำหนดปัญหาหรืองานแบบปลายเปิด เพื่อให้นักเรียนเป็นผู้สร้างคำตอบด้วยการแสดงความคิดสร้างสรรค์ หรือผลิตผลงานด้วยตนเอง
4. เน้นทักษะการคิดที่ซับซ้อนและแสดงออกอย่างชัดเจน (Complex Thinking skill) นักเรียนมีโอกาสสร้าง ผลิตผลงานจากการทำงาน มีการแก้ปัญหาจากการเรียนรู้
5. เป็นการเรียนอย่างมีความหมายต่อผู้เรียนโดยให้นักเรียนมีส่วนร่วมในการตัดสินใจว่าจะประเมินตนเองที่ไหน เรื่องอะไร อย่างไร เน้นการประเมินตนเอง (Self-Evaluation) เพื่อสะท้อนตัวผู้เรียน (Self-Reflection) เป็นการสะท้อนให้เห็นสภาพปัจจุบันของนักเรียน
6. นำผลการประเมินที่ได้สู่กระบวนการเรียนการสอนอย่างต่อเนื่อง เพื่อให้เกิดการพัฒนาในตัวผู้เรียนอย่างเต็มตามศักยภาพ
7. ผู้ประเมินควรมีหลาย ๆ คนมีการประชุมระหว่างกลุ่มผู้ประเมินอย่างหลากหลายในการประเมิน ข้อมูลได้จากหลายแหล่งโดยมีเครื่องมือเก็บรวบรวมข้อมูลหลายประเภทที่เหมาะสม เพื่อการประเมินภาพรวมของผู้เรียน และสมรรถภาพของเด็ก

เทคนิควิธีการวัดผลตามแนวของการปฏิบัติจริง มี 2 ลักษณะ คือ ประเมินจากแฟ้มสะสมงาน (Portfolio) และวัดด้วยแบบทดสอบความสามารถจริง (Authentic test) โดยกำหนดเกณฑ์การให้คะแนนแบบรูบรีค (Rubric Scoring) ในการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้การประเมินสภาพจริงด้วยแบบทดสอบความสามารถจริง (Authentic test) โดยใช้ข้อคำถามปลายเปิด

แบบทดสอบวัดความสามารถจริง (Authentic test)

ส.วาสนา ประवालพฤษ (2539 : 37) ได้กล่าวถึงแบบทดสอบวัดความสามารถจริง ไว้ดังนี้

ลักษณะของข้อสอบจะเป็นข้อคำถามที่เป็นปัญหาให้ผู้เรียนคิด และเขียนคำตอบเอง โดยปัญหาจะต้องมีความหมายต่อผู้เรียน มีความสำคัญเพียงพอ และเลียนแบบสภาพจริงในชีวิตของผู้เรียน เหมาะสมกับวัยและสภาพชีวิตของเขา แบบทดสอบทั้งฉบับจะต้องครอบคลุม และเป็นตัวแทนของเนื้อหาวิชาที่กำหนดไว้ในหลักสูตร ในการสอบผู้ตอบจะต้องใช้ความรู้ความสามารถหลายด้าน ต้องคิดอย่างลึกซึ้ง มีวิธีการและขั้นตอนที่ชัดเจนและเหมาะสม คำตอบถูกมีหลายคำตอบ วิธีการหาคำตอบมีหลายวิธี การวัดจะสะท้อนให้เห็นถึง การใช้ปัญญา ความรู้สึก และความสามารถในการปฏิบัติของผู้สอบ ให้คะแนนตามเกณฑ์ความสมบูรณ์ของคำตอบ หรือขั้นตอนที่ชัดเจนตามเกณฑ์รูบรีค (Rubric) ที่กำหนดไว้

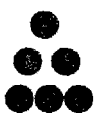
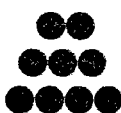
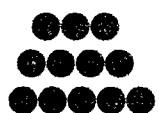
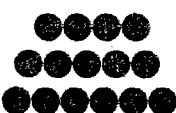
ขั้นตอนการสร้างแบบทดสอบ

1. ศึกษาหลักสูตร จัดประเด็นของหลักสูตรว่าต้องการให้ผู้เรียนเรียนรู้อะไร
2. วิเคราะห์ความรู้ความสามารถและทักษะที่สัมพันธ์กับหลักสูตรและเนื้อหาวิชา ความสามารถและเป้าประสงค์ของหลักสูตรและกระบวนการคิดที่จำเป็นในวิชานั้น
3. เขียนข้อสอบที่จะสะท้อนคุณลักษณะและกระบวนการคิดให้ครอบคลุมเนื้อหา
4. กำหนดเกณฑ์การให้คะแนนตามคุณลักษณะและกระบวนการคิด

ข้อ(00)

ทฤษฎีจำนวน เป็นการประเมินความสามารถในการใช้ความรู้พื้นฐานทางจำนวนและตัวเลขไปใช้ในการแก้ปัญหา ยลดาบอกกับพี่ชายของเธอว่า วันนี้อาจารย์ให้ฉันแบ่งแท่งไม้ ดังนี้ ถ้าฉันแบ่งแท่งไม้กลุ่มละ 2 แท่ง จะเหลือแท่งไม้ 1 แท่ง แบ่งกลุ่มละ 3 แท่ง จะเหลือแท่งไม้ 1 แท่งและถ้าแบ่งกลุ่มละ 4 แท่ง จะเหลือแท่งไม้ 1 แท่ง เช่นเดียวกัน จงหาว่ายลดา มีแท่งไม้ทั้งหมดกี่แท่ง แสดงการแก้ปัญหา คำตอบ

ข้อ(000)

รูปแบบ เป็นการประเมินทักษะการให้เหตุผลของนักเรียน เพื่อบอกหลักเกณฑ์ที่ใช้ในการสร้างรูปภาพ รูปถัดไปจากรูปแบบที่กำหนดให้ จากรูปแบบข้างล่างนี้ จงเขียนรูปถัดไป พร้อมบอกหลักเกณฑ์ในการสร้าง     ก. รูปถัดไปคือ ข. หลักเกณฑ์ในการสร้างรูป คือ

4. เกณฑ์การให้คะแนน

เครื่องมือที่ใช้ในการประเมินการปฏิบัติงานของนักเรียนเรียกว่า "รูบรีค" (Rubric) ซึ่งจะกำหนดมาตรวัด (Scale) และรายการของคุณลักษณะ ที่บรรยายถึงความสามารถในการแสดงออก ของแต่ละจุดในมาตรวัดไว้อย่างชัดเจน การให้คะแนนรูบรีค มี 2 แบบ (กรมวิชาการ.2539 : 54-59) คือ

1. การให้คะแนนเป็นภาพรวม (Holistic score) คือ การให้คะแนนงานชิ้นใดชิ้นหนึ่งโดยคุณภาพรวมของชิ้นงานว่ามีความเข้าใจในความคิดรวบยอด การสื่อความหมาย กระบวนการที่ใช้และผลงานเป็นอย่างไร แล้วเขียนอธิบายคุณภาพของงานหรือความสำเร็จของงานเป็นชิ้น ๆ โดยอาจจะแบ่งระดับคุณภาพตั้งแต่ 0-4 หรือ 0-8 สำหรับในขั้นต้นการให้คะแนนรูบรีค อาจจะแบ่งวิธีการให้คะแนนหลายวิธี เช่น

วิธีที่ 1 แบ่งงานตามคุณภาพเป็น 3 กอง คือ

กองที่ 1 ได้แก่ งานที่มีคุณภาพเป็นพิเศษและเขียนอธิบายลักษณะของงานที่มีคุณภาพเป็นพิเศษ

กองที่ 2 ได้แก่ งานที่ยอมรับได้และเขียนอธิบายลักษณะของงานที่ยอมรับได้

กองที่ 3 ได้แก่ งานที่ยอมรับได้น้อย หรือยอมรับไม่ได้ และเขียนอธิบายลักษณะของงานที่ยอมรับได้น้อยจากนั้นก็นำงานแต่ละกองมาให้คะแนนเป็น 2 ระดับ คือ

งานกองที่ 1 จะให้คะแนน 6 หรือ 5

งานกองที่ 2 จะให้คะแนน 4 หรือ 3

งานกองที่ 3 จะให้คะแนน 2 หรือ 1

วิธีที่ 2 กำหนดตามระดับความผิดพลาด คือพิจารณาความบกพร่อง จากคำตอบว่ามีมากน้อยเพียงใด โดยจะหักจากระดับคะแนนสูงสุดลงมาทีละระดับ ดังนี้

คะแนน 4 หมายถึง คำตอบถูกต้องแสดงผลถูกต้อง แนวคิดชัดเจน

3 หมายถึง คำตอบถูกต้อง เหตุผลถูกต้อง แต่มีข้อผิดพลาดเล็กน้อย

2 หมายถึง เหตุผลหรือการคำนวณผิดพลาด แต่มีแนวทางที่จะนำไปสู่คำตอบ

1 หมายถึง แสดงวิธีคิดเล็กน้อยแต่ไม่ได้คำตอบ

0 หมายถึง ไม่ตอบหรือตอบไม่ถูกเลย

วิธีที่ 3 กำหนดระดับการยอมรับและคำอธิบาย เช่น

รูบรีค (Rubric) ของความสามารถเข้าใจเนื้อหาสาระ เขียนได้เป็น 4 ระดับ ดังนี้

- 4 หมายถึง การสาธิตหรือแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์ ครบถ้วนถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนด รวมทั้งเสนอแนวคิดใหม่ ที่แสดงถึงความเข้าใจอย่างลึกซึ้งถึงกฎเกณฑ์ หรือลักษณะของข้อมูล
- 3 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์ ครบถ้วน ถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนด
- 2 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจที่ไม่สมบูรณ์ ครบถ้วน ถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนดในบางส่วน
- 1 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนดได้น้อยมาก และเข้าใจไม่ถูกต้องบางส่วน
- 0 หมายถึง ไม่แสดงความคิดเห็นใด ๆ

2. การให้คะแนนแบบแยกองค์ประกอบ (Analytic score) เพื่อให้การมองคุณภาพงานหรือความสามารถของนักเรียนได้อย่างชัดเจน จึงได้มีการแยกองค์ประกอบของการให้คะแนน และการอธิบายคุณภาพของงานในแต่ละองค์ประกอบเป็นระดับ โดยทั่วไปแล้วจะมีการแยกองค์ประกอบของงานเป็น 4 ด้าน คือ

2.1 ความเข้าใจในความคิดรวบยอด ข้อเท็จจริง เป็นการแสดงให้เห็นว่านักเรียนเข้าใจในความคิดรวบยอด หลักการในการแก้ปัญหาที่ถามอย่างกระจ่างชัด

2.2 การสื่อความหมาย สื่อสาร คือ ความสามารถในการอธิบาย นำเสนอ การบรรยาย เหตุผล แนวคิด ให้ผู้อื่นเข้าใจได้ดี มีความคิดสร้างสรรค์

2.3 การใช้กระบวนการและยุทธวิธี สามารถเลือกใช้ยุทธวิธี กระบวนการนำไปสู่ความสำเร็จอย่างมีประสิทธิภาพ

2.4 ผลสำเร็จของงาน ความถูกต้องแม่นยำในผลสำเร็จของงาน หรืออธิบายที่มาและตรวจสอบผลงาน

ตัวอย่าง การให้คะแนนแบบแยกองค์ประกอบ (Analytic score) (Reys. 1992 : 313)

เรย์ (Reys. 1992 : 313) ได้กำหนดระดับของความสามารถในการแก้ปัญหา โดยที่แต่ละขั้นตอนของกระบวนการแก้ปัญหา จะให้คะแนนตั้งแต่ 0-2 คะแนน ตามรายละเอียดดังนี้

1. ความเข้าใจในปัญหา

0 : ไม่เข้าใจในปัญหาเลย

1 : เข้าใจปัญหามบางส่วน หรือแปลความหมายบางส่วนคลาดเคลื่อน

2 : เข้าใจปัญหาได้ดี ครบถ้วนสมบูรณ์

2. การวางแผนการแก้ปัญหา

0 : ไม่พยายาม หรือวางแผนได้ไม่เหมาะสมทั้งหมด

1 : วางแผนถูกต้องบางส่วน ในส่วนที่แปลความหมายได้ถูกต้อง

2 : วางแผนเพื่อนำไปสู่การแก้ปัญหาได้ถูกต้องทั้งหมด

3. คำตอบ

0 : ไม่ตอบ หรือตอบผิดในส่วนที่วางแผนไม่เหมาะสม

1 : คัดลอกผิดพลาด คำนวณผิดพลาด ตอบบางส่วนสำหรับปัญหาที่มีหลายคำตอบ

2 : ตอบได้ถูกต้อง และใช้ภาษาได้ถูกต้อง

นอกจากนี้ ส.ว.ส.นา ประมวลพฤษ (2539 : 38-40) ได้เสนอการกำหนดเกณฑ์การให้คะแนนแบบรูปรีด (Rubric) เพื่อประเมินแฟ้มสะสมงานหรือแบบทดสอบ ไว้ 5 วิธีดังต่อไปนี้

วิธีที่ 1 แยกประเด็นพิจารณาออกเป็นประเด็นย่อยแล้วทำเป็นตารางพิจารณาความถูกต้องในแต่ละประเด็น กำหนดระดับของคะแนนตามจำนวนที่ปฏิบัติถูกต้อง ในประเด็นเหล่านั้น

ตัวอย่าง กำหนดให้นักเรียนศึกษาว่ากระดาษทิชชู 3 ยี่ห้อ ยี่ห้อไหนจะซับน้ำได้ดีที่สุด โดยใช้อุปกรณ์การทดลองประกอบด้วย หลอดแก้ว ถาด หลอดหยด และตาชั่ง

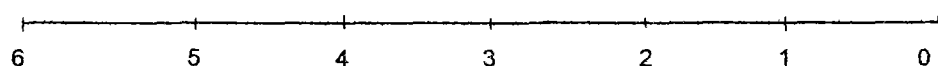
เกณฑ์การให้คะแนนจะพิจารณาวิธีการ การทำให้ม้วนตัว การพิจารณาผล การชั่งน้ำหนัก และการสรุปผล

คะแนน	วิธีการ	การทำให้อึดตัว	การพิจารณาผล	การชั่ง	ผล
4	ถูก	ถูก	ถูก	ถูก	ถูก
3	ถูก	ถูก	ถูก	ผิด	ถูก/ผิด
2	ถูก	ถูก	คลาดเคลื่อน	ผิด	ถูก/ผิด
1	ถูก	ผิด	คลาดเคลื่อน	ผิด	ถูก/ผิด
0 ไม่ได้ปฏิบัติเลย/ปฏิบัติผิดหมด					

วิธีที่ 2 กำหนดระดับความสมบูรณ์ ตามเส้นแสดงความต่อเนื่องของความสามารถ
(Continuous ability)

ตัวอย่าง แบบทดสอบวัดความสามารถในการแก้ปัญหา

- ผลงานถูกต้อง
- แสดงวิธีคิดมากกว่าวิธีการหาคำตอบ
- บอกได้ว่าทำไมคำตอบจึงถูก
- บอกแนวทางการแก้ปัญหาแนวทางอื่น ๆ ได้



- บางตอนหรือบางส่วนของคำตอบไม่ถูกต้อง
- ก่อนข้างจะบอกได้ว่าคำตอบนั้นถูกต้อง
- ทราบวิธีการแต่แก้ปัญหาไม่สำเร็จ
- งานไม่สำเร็จ
- แสดงความพยายามแก้ปัญหา
- ไม่สามารถจะบอกได้ว่าทำไมจึงแก้ปัญหาแนวนี้

จากลักษณะของเกณฑ์การให้คะแนนดังกล่าว ผู้วิจัยเลือกใช้เกณฑ์การให้คะแนนแบบการประเมินรวม (Holistic score) ประเภทกำหนดตามระดับความผิดพลาด

5. หลักฐานความเที่ยงตรงเชิงโครงสร้าง

หลักฐานความเที่ยงตรงเชิงโครงสร้าง หมายถึง ข้อมูลที่แสดงว่าเครื่องมือนั้นสามารถวัดได้ตรงตามลักษณะของโครงสร้างที่ต้องการวัดได้ชัดเจน ซึ่งแสดงโดยใช้ค่าสถิติต่าง ๆ หรือการทดลองจัดกระทำให้ประจักษ์ชัด

การคำนวณค่าความเที่ยงตรงเชิงโครงสร้าง

ในการคำนวณค่าความเที่ยงตรงเชิงโครงสร้าง กระทำได้หลายวิธี ดังนี้

1. ให้ผู้เชี่ยวชาญวิเคราะห์ (Expert) เพื่อตรวจสอบว่าข้อสอบแต่ละข้อวัดได้ตรงตามพฤติกรรมในตารางวิเคราะห์หลักสูตรหรือไม่

2. ใช้เทคนิคกลุ่มที่รู้จัก (Known Group Technique) แอลเลนและเยน (Allen and Yen. 1979 : 108) ได้เสนอแนวคิดนี้ไว้ว่า ถ้าใช้กลุ่มตัวอย่างที่เราทราบว่ามีความลักษณะอย่างนั้นอยู่แล้วกับกลุ่มที่ไม่มี ทำแบบทดสอบฉบับเดียวกัน แล้วนำคะแนนที่ได้มาเปรียบเทียบกันโดยวิธีทางสถิติ ถ้าคะแนนทั้งสองกลุ่มแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ก็แสดงว่าเครื่องมือที่มีความเที่ยงตรงเชิงโครงสร้างสูง

3. หาค่าสหสัมพันธ์กับแบบทดสอบอื่น (Correlation with Other Tests) อนาสตาซี (บุญเชิด ภิญญอนันตพงษ์. 2521 : 345-355 อ้างอิงมาจาก Anatsasi. 1968 : 115) ได้เสนอว่า สหสัมพันธ์ระหว่างแบบทดสอบที่มีอยู่แล้วกับแบบทดสอบใหม่ ซึ่งวัดพฤติกรรมด้านเดียวกัน จะสามารถบอกความเที่ยงตรงเชิงโครงสร้างได้ โดยแบบทดสอบที่มีอยู่แล้วนั้นจะต้องเป็นแบบทดสอบที่เป็นมาตรฐาน นั่นคือเมื่อนำเครื่องมือที่จะหาความเที่ยงตรงเชิงโครงสร้าง กับเครื่องมือที่วัดคุณลักษณะเดียวกันที่เป็นมาตรฐาน ไปสอบกับกลุ่มตัวอย่างกลุ่มเดียวกัน แล้วนำคะแนนที่ได้มาหาค่าสหสัมพันธ์โดยใช้สูตรของเพียร์สัน (r_{xy}) ค่าสหสัมพันธ์ที่ได้คือ ค่าความเที่ยงตรงเชิงโครงสร้าง

4. การวิเคราะห์องค์ประกอบ (Factor Analysis) แอลเลนและเยน (Allen and Yen. 1979 : 111) กล่าวว่า การวิเคราะห์องค์ประกอบ เป็นการวิเคราะห์ตัวประกอบโดยอาศัยวิธีการทางสถิติ เพื่อวิเคราะห์สหสัมพันธ์ระหว่าง กลุ่มของตัวแปรในรูปของเมทริกซ์สหสัมพันธ์ และอธิบายสหสัมพันธ์เหล่านั้นในรูปของการลดจำนวนตัวแปรต่าง ๆ ให้น้อยลง ซึ่งเรียกว่า ตัวประกอบ จนกระทั่งเหลือแต่ตัวประกอบร่วมที่สำคัญ แบบทดสอบที่ได้รับอิทธิพลจากตัวประกอบโดยอ้อมมีค่านำหนักตัวประกอบสูง (high factor loading) และค่านำหนักตัวประกอบแต่ละตัวจะแสดงถึง ค่าความเที่ยงตรงเชิงโครงสร้าง ถ้าค่านำหนักในตัวประกอบโดยของแบบทดสอบมีค่าใกล้เคียงกันมาก แสดงว่าแบบทดสอบชุดนี้วัดสิ่งเดียวกัน นั่นคือแบบทดสอบมีความเที่ยงตรงเชิงโครงสร้างสูง

5. เมทริกซ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod Matrix) เป็นการค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ ที่พัฒนาโดย แคมป์เบลล์และฟิเชอร์ (Allen and Yen. 1979 : 109 -110 ; citing Cambell and Fiske.1959) เมื่อแบบทดสอบที่สร้างขึ้นมีวิธีวัด 2 วิธี หรือมากกว่า และคุณลักษณะที่วัดมี 2 ลักษณะหรือมากกว่า โดยพิจารณาค่าความเที่ยงตรงเป็น 2 ลักษณะ คือ ความเที่ยงตรงเชิงเหมือน (Convergent Validity) และ ความเที่ยงตรงเชิงจำแนก (Discriminant Validity) ดังแสดงในตาราง 2

ตาราง 2 เมทริกซ์ หลายลักษณะ-หลายวิธี เมื่อวัดคุณลักษณะ A และ B ด้วยวิธีวัด 1 และ 2

คุณลักษณะ/วิธีวัด	A ₁	B ₁	A ₂	B ₂
A ₁	สูง	ต่ำ	สูง	ต่ำ
B ₁	ต่ำ	สูง	ต่ำ	สูง
A ₂	สูง	ต่ำ	สูง	ต่ำ
B ₂	ต่ำ	สูง	ต่ำ	สูง

จากตาราง ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ ที่ใช้วิธี หาค่าความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) ประกอบด้วย 2 ลักษณะ คือ ความเที่ยงตรงเชิงเหมือน (Convergent Validity) ซึ่งเป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จาก การวัดลักษณะเดียวกันด้วยวิธีวัดต่างกัน ($r(A_1A_2)$, $r(B_1B_2)$) ค่าสหสัมพันธ์จะสูง และความเที่ยงตรงเชิงจำแนก (Discriminant Validity)) ซึ่งเป็นค่า

สัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดลักษณะต่างกัน ด้วยวิธีวัดวิธีเดียวกันหรือต่างวิธีกัน ($r(A_1 B_1), r(A_2 B_2), r(A_1 B_2), r(A_2 B_1)$) ค่าสหสัมพันธ์จะต่ำ และมีค่าต่ำกว่าความเที่ยงตรงเชิงเหมือน (Convergent Validity) และจากตาราง ค่าสหสัมพันธ์ที่คำนวณได้ ประกอบด้วย

1. สัมประสิทธิ์สหสัมพันธ์ในแนวทะแยงมุมหลัก ของเมทริกซ์ ($r(A_1 A_1), r(B_1 B_1), r(A_2 A_2), r(B_2 B_2)$) คือสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบที่ได้จากการสอบซ้ำ หรือคู่ขนาน ควรจะมีค่าสูงสุด
2. สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสอบที่ได้ จากการวัดคุณลักษณะเดียวกัน ด้วยวิธีวัดต่างวิธีกัน ($r(A_1 A_2), r(B_1 B_2)$) ควรจะมีค่าสูง
3. สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสอบที่ได้จากการวัดคุณลักษณะต่างกัน ด้วยวิธีวัดวิธีเดียวกัน ($r(A_1 B_1), r(A_2 B_2)$) ควรจะมีค่าต่ำ
4. สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสอบที่ได้ จากการวัดคุณลักษณะต่างกัน ด้วยวิธีวัดต่างวิธีกัน ($r(A_1 B_2), r(A_2 B_1)$) ควรจะมีค่าต่ำสุด

แต่ก็อาจเป็นไปได้ที่ลักษณะต่างกันแต่วัดด้วยวิธีเดียวกัน แล้วพบว่า ผลการวัดมีสหสัมพันธ์กันสูง ลักษณะเช่นนี้เรียกว่าเกิดผลที่เหมือนกันของวิธี (Method/halo effect) (Marsh, Smit & Barnes. 1983.) ซึ่งถือว่าเป็นความคลาดเคลื่อนที่เกิดขึ้น โดยมีได้ปรารถนาที่จะให้ค่าสหสัมพันธ์สูง เช่นในการประเมินคุณภาพการสอนของอาจารย์นั้นจะต้องวัดหลายลักษณะ เช่น เตรียมการสอน การบรรยาย การใช้อุปกรณ์การสอน ในการประเมินลักษณะต่าง ๆ เหล่านี้ใช้นิสิตประเมินทั้ง 3 ลักษณะ ถ้าเป็นอาจารย์ที่นิสัยนิยมยกย่องมากในเรื่องอื่น ๆ ที่ไม่เกี่ยวกับการสอน ก็อาจจะทำให้ค่าสหสัมพันธ์ของลักษณะทั้งสามนี้สูงก็ได้ และอาจเป็นไปได้อีกเช่นกันว่าลักษณะต่างกันและวิธีวัดก็ต่างกันแต่ผลการวัดมีค่าสหสัมพันธ์กันสูง เหตุการณ์เช่นนี้ถือว่าเป็นกรณีที่น่าวิตกอย่างยิ่งกับความเที่ยงตรงเชิงจำแนก สิ่งนี้อาจเกิดขึ้นได้เนื่องจากโครงสร้างของสิ่งที่วัดทั้งสองนั้นมีได้แยกกันอย่างเด่นชัดนัก

ในกรณีที่คุณลักษณะที่ต้องการวัดมีมากกว่า 2 ลักษณะ และวิธีวัดมีมากกว่า 2 วิธี เมทริกซ์สหสัมพันธ์จะมีขนาดใหญ่ขึ้น และมีความซับซ้อนมากขึ้น แต่การพิจารณาค่าความเที่ยงตรงของแบบทดสอบยังคงใช้หลักเดิม คือ สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบที่ได้จากการสอบซ้ำ หรือคู่ขนาน ควรจะมีค่าสูงสุด สัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดคุณลักษณะเดียวกัน ควรจะมีค่าสูงกว่า สัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดคุณลักษณะต่างกัน

จากค่าความสัมพันธ์หลายลักษณะ-หลายวิธี สามารถสรุปเป็นเกณฑ์ในการวิเคราะห์ความเที่ยงตรงของแบบทดสอบได้ดังนี้ (Marsh, Smit and Barnes. 1983 ; Centra.1971)

1. ค่าความเชื่อมั่นของแบบทดสอบจะสูงกว่าค่าความเที่ยงตรงเหมือน
2. ค่าความเที่ยงตรงเชิงเหมือนจะมีค่ามากกว่าค่าความเที่ยงตรงเชิงจำแนก
3. ค่าความเที่ยงตรงเชิงเหมือนจะมีค่าสูง โดยปกติจะมากกว่าศูนย์อย่างมีนัยสำคัญทางสถิติ
4. ค่าความเที่ยงตรงเชิงจำแนกจะมีค่าต่ำและมักเป็นศูนย์

การหาความเที่ยงตรงเชิงโครงสร้างตามสูตรของ ไฮเซน และบรอนสเตอร์ด

สุทธิพงศ์ บุญผดุง (2541 : 10-13) ได้ศึกษาสูตรการหาค่าความเที่ยงตรงเชิงโครงสร้างของ ไฮเซน และบรอนสเตอร์ด (Heise and Bohmstedt. 1970 : 109, 111-113) ซึ่งไฮเซน และบรอนสเตอร์ด ได้เสนอวิธีหา ค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ โดยใช้ทฤษฎีตัวประกอบร่วม (Common Factor Theory) ในรูปของการหาสหสัมพันธ์(Correlation)ระหว่างคะแนนจริงของคุณลักษณะที่ได้ T_k กับคะแนนองค์ประกอบ S_k เมื่อคะแนนจริงของคุณลักษณะ (Traits) สัมพันธ์กันแบบออร์โธโกนอล (Orthogonal : $\rho_{T_i T_j} = 0$) นั่นคือ แต่ละคุณลักษณะ (T_1, T_2) ไม่มีความสัมพันธ์กัน และการหาน้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k สามารถหาได้จาก การวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis) จากการหมุนแกนแบบมุมฉาก ด้วยวิธีแวนริแมกซ์ (Varimax) และสกัดองค์ประกอบด้วยวิธี PC (Principal Component Analysis) โดยใช้โปรแกรม SPSS

สูตรสำหรับหาค่าความเที่ยงตรงเชิงโครงสร้าง ที่สุทธิพงศ์ บุญผดุง (2541 : 10-13) ได้ศึกษาจากสูตรของไฮเซน และบรอนสเตอร์ด (Heise and Bohmstedt. 1970 : 109, 111-113) มีดังนี้

$$\rho_{T_i S_k} = \frac{\sum_{i=1}^n S_i f_{ik}}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n \text{Cov}(x_i, x_j)}}$$

เมื่อ	$\rho_{T_i S_k}$	แทน	ค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ
	S_i	แทน	ค่าเบี่ยงเบนมาตรฐานของคำถามข้อที่ i
	S_{ik}	แทน	น้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k
	$\text{Cov}(x_i, x_j)$	แทน	ค่าความแปรปรวนร่วมกันระหว่างคำถามข้อที่ i กับคำถามข้อที่ j

จากวิธีการคำนวณค่าความเที่ยงตรงเชิงโครงสร้างดังกล่าว ในการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้วิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) ที่พัฒนาโดย แคมป์เบลล์และฟิสก์ (Allen and Yen. 1979 : 109 -110 ; citing Cambell and Fiske.1959) และ คำนวณด้วยสูตร $\rho_{T_i S_k}$ ไฮเซน และบรอนสเตอร์ด (Heise and Bohmstedt. 1970 : 109, 111-113)

6. ความเชื่อมั่นของแบบทดสอบ

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

1. การวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ ของการวัดหลาย ๆ ครั้ง จากการสอบซ้ำด้วยแบบทดสอบฟอร์มเดียวกัน หรือแบบทดสอบฟอร์มที่คู่ขนานกัน ซึ่งมีค่าเท่ากับ ความแปรปรวนของคะแนนจริงหารด้วยความแปรปรวนของคะแนนที่สอบได้ โดยเรียกสัมประสิทธิ์สหสัมพันธ์ ที่คำนวณจากคะแนนที่ได้จากการสอบด้วยข้อสอบฉบับเดียวกันสองครั้งในเวลาที่แตกต่างกันว่า สัมประสิทธิ์ของความคงตัว (Coefficient of Stability) และเรียกสัมประสิทธิ์สหสัมพันธ์ ที่คำนวณจากคะแนนที่ได้จากการสอบด้วยข้อสอบคู่ขนาน (Parallel Test) หรือแบบสลับฟอร์ม (Alternate Form) ที่สอบติดต่อกันทันที หรือสอบทิ้งช่วงเวลาก็ได้ ว่า สัมประสิทธิ์ของความสมมูล (Coefficient of Equivalent)

2. ใช้แบบทดสอบเพียงฉบับเดียวและสอบเพียงครั้งเดียว แล้วคำนวณหาสัมประสิทธิ์ความสอดคล้องภายในของแบบทดสอบ (Coefficient of Internal Consistency) จากความแปรปรวนของคะแนนสอบภายในฉบับจากส่วนย่อยแต่ละตอนและส่วนรวมทั้งฉบับ

นักทฤษฎีทางการวัดผลได้เสนอเทคนิค ในการประมาณค่าความเชื่อมั่นที่เชื่อถือได้ไว้หลายเทคนิค โดยมีข้อสันนิษฐานเบื้องต้น (presumptions) ว่า แบบทดสอบฉบับรวมสามารถแบ่งเป็นส่วน ๆ เช่น สองส่วน สามส่วน สี่ส่วน หรือหลาย ๆ ส่วน และเมื่อใช้ระดับของความคู่ขนานของการวัด (degree of measurement parallelism) ในแต่ละส่วนเป็นเกณฑ์แล้ว การประมาณค่าความเชื่อมั่นที่ได้จากการสอบเพียงครั้งเดียวด้วยข้อสอบเพียงฉบับเดียว จะจัดกลุ่มตามข้อตกลงของระดับความคู่ขนานได้ 3 กลุ่ม ดังนี้

แบบจำลองความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel Parts)

เป็นการประมาณค่าความเชื่อมั่นจากแบบทดสอบแต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิม ที่มีข้อตกลงอย่างเคร่งครัด 6 ข้อ (บุญเชิด ภิญญอนันตพงษ์ . 2536 : 9-10) คือ

1. มีความเป็นเอกพันธ์ในเนื้อหา หรือวัดคุณลักษณะเดียวกัน
2. มีคะแนนจริงเท่ากัน ($T_{11} = T_{12} = T_{13} = \dots$) และมีความแปรปรวนคลาดเคลื่อนเท่ากัน ($S_{E1} = S_{E2} = S_{E3} = \dots$)
3. มีคะแนนสอบ (X) เฉลี่ยเท่ากัน ($\mu_1 = \mu_2 = \mu_3 = \dots$)
4. มีความแปรปรวนของคะแนนสอบ (X) เท่ากัน ($S^2_1 = S^2_2 = S^2_3 = \dots$)
5. มีความแปรปรวนร่วมของคะแนนสอบ (X) กับคะแนนสอบส่วนอื่น ๆ เท่ากัน ($S_{12} = S_{13} = S_{23} = \dots$)
6. มีความแปรปรวนร่วมของคะแนนสอบ (X) กับคะแนนเกณฑ์ภายนอกเท่ากัน ($S_{1y} = S_{2y} = S_{3y} = \dots$)

จากข้อตกลงดังกล่าว ในทางปฏิบัติแทบไม่สามารถสร้างแบบทดสอบให้แต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิมได้ จึงได้มีการพัฒนาเทคนิคที่เหมาะสมขึ้นมาใหม่ โดยผ่อนปรนเงื่อนไขเดิมมาเป็น แต่ละส่วนจำเป็นต้องมีคะแนนจริงสมมูล (Essentially Tau-Equivalent Parts)

นักทฤษฎีการทดสอบที่มีชื่อเสียงได้แก่ สเปียร์แมน (Spearman. 1910) และบราวน์ (Brown. 1910)

แบบจำลองคะแนนจริงจำเป็นต้องสมมูล(Essentially Tau-Equivalent Parts)

วิธีนี้เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบ โดยผ่อนปรนเงื่อนไข ข้อ 2,3 และ 4 ของแบบจำลองความคู่ขนานแบบมาตรฐานเดิม ให้มีความเป็นไปได้มากขึ้น (บุญเชิด ภิญโญนันตพงษ์. 2536 : 10-11) ดังนี้

- คะแนนจริงแต่ละส่วนไม่จำเป็นต้องเท่ากัน แต่ยอมให้ต่างกันได้เท่ากับความยาวที่ต่างกันในแต่ละส่วน นั่นคือ $T_{hg} = T_{hh} + C_{gh}$ เมื่อ $g = h = 1, \dots, k$ และ C_{gh} ไม่จำเป็นต้องเท่ากับศูนย์เสมอไป
- แต่ละส่วนมีคะแนนสอบ (X) เฉลี่ยต่างกัน
- ความแปรปรวนของคะแนนสอบ (X) ต่างกันเล็กน้อย

แม้ว่าแต่ละส่วนของแบบทดสอบที่แบ่งนั้นจะมีขนาดของความยาวหรือจำนวนข้อเท่ากัน แต่ในทางปฏิบัติจริงมีแบบทดสอบบางชนิดอาจต้องแบ่งส่วนให้เหมาะสมตามลักษณะของแบบทดสอบ ทำให้แต่ละส่วนมีจำนวนข้อไม่เท่ากัน ซึ่งส่งผลต่อเงื่อนไขข้อ 5 และ 6 หรือแม้ว่าแต่ละส่วนจะประกอบด้วยจำนวนข้อที่เท่ากัน แต่เมื่อนำไปสอบกับกลุ่มตัวอย่างแล้วปรากฏว่า แต่ละส่วนมีการกระจายของคะแนนไม่เท่ากัน แสดงว่าความยาวที่ทำหน้าที่ (Functional lengths) ของมันในแต่ละส่วนมีขนาดไม่เท่ากัน (Feldt. 1975 : 558) นักทฤษฎีการทดสอบจึงจำเป็นต้องนิยามความคู่ขนานขึ้นมามีอีกระดับหนึ่ง ที่มีความผ่อนปรนมากที่สุด เรียกว่าความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model)

นักทฤษฎีที่เสนอเทคนิคสำหรับการประมาณค่าความเชื่อมั่นตามกลุ่มนี้ ได้แก่ ฟลานาแกน(Flanagan. 1930) รูลอน (Rulon. 1930) กัตต์แมน(Guttman. 1945) เฟลด์และเบรนนอน(Feldt and Brennan. 1989) ครอนบาค (Cronbach. 1951)

ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric model)

เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบที่ผ่อนปรนเงื่อนไขต่างๆ เกือบทั้งหมดโดยคงไว้เฉพาะเงื่อนไขข้อ 1 ที่ว่า แต่ละส่วนของแบบทดสอบต้องมีเนื้อหาที่เป็นเอกพันธ์ หรือวัดคุณลักษณะเดียวกัน (Kristof. 1974 : 492)

ลักษณะสำคัญของความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ คือการแบ่งแบบทดสอบออกเป็น ส่วน ๆ ที่มีขนาดความยาวไม่เท่ากัน หรือมีขนาดความยาวเท่ากันแต่มีการกระจายของคะแนนในแต่ละส่วนแตกต่างกันมาก การประมาณค่าความเชื่อมั่นนอกจากจะใช้สถิติขั้นสูงวิเคราะห์ด้วยเครื่องคอมพิวเตอร์แล้ว นักทฤษฎีการวัดยังได้เสนอเทคนิคที่ใช้เครื่องคำนวณแบบธรรมดาไว้หลายเทคนิค ดังนี้

ฮอริส(บุญเชิด ภิญโญนันตพงษ์. 2533 : 29 ; อ้างอิงจาก Horst. 1951: 368-371) เสนอสูตรของการแบ่งแบบทดสอบเป็นสองส่วน มีความยาวไม่เท่ากัน ดังนี้

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

r_u แทน ค่าความเชื่อมั่นของแบบทดสอบ

r แทน สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของแบบทดสอบส่วนย่อยทั้งสองส่วน

λ_1, λ_2 แทน สัดส่วนของจำนวนข้อแต่ละส่วนย่อยจากจำนวนข้อทั้งหมด

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

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

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

เฟลด์ต์(บุญเชิด ภิญโญนนตพงษ์. 2533 : 30 ; อ้างอิงจาก Feldt. 1975 : 557-561) ได้ศึกษาค่าความเชื่อมั่นของแบบทดสอบที่แบ่งสามส่วนไม่เท่ากันของคริสทอฟ ให้สามารถวิเคราะห์กับแบบทดสอบที่แบ่งสองส่วนไม่เท่ากัน โดยเพิ่มข้อตกลงอีกหนึ่งข้อ คือ ความแปรปรวนของคะแนนความคลาดเคลื่อน ต้องเป็นไปตามทฤษฎีการวัดแบบมาตรฐานเดิม และไม่จำเป็นต้องทราบจำนวนข้อของแบบทดสอบ ดังนี้

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

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

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

เมื่อ

r_{tt} แทน ค่าความเชื่อมั่นของแบบทดสอบ
 λ_1, λ_2 แทน สัดส่วนของจำนวนข้อแต่ละส่วนย่อย
 s_{x1}^2 แทน ความแปรปรวนของคะแนนย่อยส่วนที่ 1
 s_{x2}^2 แทน ความแปรปรวนของคะแนนย่อยส่วนที่ 2
 s_{xij} แทน ความแปรปรวนร่วม

ราชู (บุญเขต วิทยุโณนันทพงษ์. 2533 : 30 ; อ้างอิงจาก Raju. 1977 : 549-565) ได้พัฒนาสูตรการประมาณค่าความเชื่อมั่น เมื่อแบ่งแบบทดสอบออกเป็นหลาย ๆ ส่วน แต่ละส่วนมีความยาวไม่เท่ากัน เรียกว่า สัมประสิทธิ์เบต้าเค (Coefficient- β_k) แต่ถ้าแต่ละส่วนมีจำนวนข้อเท่ากันแล้ว ค่าสัมประสิทธิ์เบต้าเค จะเท่ากับสัมประสิทธิ์แอลฟาของครอนบัก ดังนี้

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

เมื่อ	β_k	แทน ค่าความเชื่อมั่นของแบบทดสอบ
	λ_i	แทน สัดส่วนของจำนวนข้อแต่ละส่วนย่อย
	s_i^2	แทน ความแปรปรวนของคะแนนในแต่ละส่วน
	s_x^2	แทน ความแปรปรวนของคะแนนทั้งฉบับ

กิล เมอร์และเฟลด์ต์ (บุญเขต วิทยุโณนันทพงษ์. 2537 : 26-30 ; อ้างอิงจาก Gilmer and Feldt. 1983 : 101-105) ได้เสนอสูตรการประมาณค่าความเชื่อมั่น โดยพยายามลดจำนวนไม่ทราบค่า $K(K+1)/2 + 1$ ให้เหลือเท่ากับความสัมพันธ์ของตัวแปร $2K+1$ แล้วแก้สมการตัวไม่ทราบค่า $2K+1$ ซึ่งมีสูตรวิเคราะห์ 2 สูตร คือ r_{F1} และ r_{F2} สามารถใช้คำนวณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งเป็นหลาย ๆ ส่วน ($K > 3$)

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

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

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

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

r_{F2}	แทน ค่าความเชื่อมั่นของแบบทดสอบ
s_{KL}	แทน ความแปรปรวนร่วมระหว่างส่วนย่อย K และ L
L	แทน แถวที่ผลรวมของความแปรปรวนร่วมมีค่าสูงสุด

เลี้ยว(บุญเชิด วิทยุโณนันทพงษ์. 2537 : 33-37 ; อ้างอิงจาก Liou.1989 : 153-163) ได้ทำการวิจัย และเสนอสัมประสิทธิ์ความเชื่อมั่นขึ้นมาใหม่ สองสัมประสิทธิ์ คือ r_{L1} และ r_{L2} ซึ่งมีข้อดีที่สามารถคำนวณด้วยเครื่องคำนวณธรรมดา และยังสามารถพิสูจน์ได้ทั้งเชิงทฤษฎีและจำนวนตัวเลข ว่ามีค่าเท่ากับสัมประสิทธิ์ r_{Lk} และ r_{LF} ของคริสทอฟ(Kristof.1974) และของเฟลด์ต์(Feldt.1975) เมื่อ K เท่ากับ สามหรือสองตอนตามลำดับ มีสูตรในการคำนวณ ดังนี้

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

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

...

$$Q_k = \frac{(\sum s_{gk})^2 - \sum s_{gk}^2}{\sum \sum_{g \neq h} s_{gh}}, g, h \neq k$$

- เมื่อ r_{L1} แทน ค่าความเชื่อมั่นของแบบทดสอบ
 S_{gh} แทน ความแปรปรวนร่วมระหว่างส่วนย่อยของแบบทดสอบ
 S_g^2 แทน ความแปรปรวนของคะแนนในแต่ละตอน
 S_x^2 แทน ความแปรปรวนของคะแนนสอบทั้งฉบับ

สูตร r_{L2} เมื่อแบ่งแบบทดสอบเป็น k ส่วน

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

$$\lambda_1 = \frac{\sum s_{1g}}{s_x^2}$$

...

$$\lambda_k = \frac{\sum s_{kg}}{s_x^2}$$

- เมื่อ r_{12} แทน ค่าความเชื่อมั่นของแบบทดสอบ
 s_g^2 แทน ความแปรปรวนของคะแนนในแต่ละตอน
 s_x^2 แทน ความแปรปรวนของคะแนนสอบทั้งฉบับ
 λ_g^2 แทน ผลรวมของความแปรปรวนร่วมระหว่างส่วนย่อยของแบบทดสอบ

นอกจากนี้ยังมีนักวิจัยอีกกลุ่มหนึ่ง ได้เสนอสูตรที่ใช้คำนวณค่าความเชื่อมั่น โดยให้ความสำคัญกับการนำค่าน้ำหนักองค์ประกอบ (Factor Loading) มาใช้ในการคำนวณค่าความเชื่อมั่นของแบบทดสอบ อย่างเช่น ไฮเซนและบรอนสเตอร์ท (Heise and Bohmstedt. 1910) ได้เสนอสูตร Ω ซึ่งเป็นสัมประสิทธิ์ค่าความเชื่อมั่นของแบบทดสอบ เมื่อกำหนดน้ำหนักความสำคัญเท่ากัน ต่อมาอัลเลน (Allen. 1974) ได้พัฒนาสูตรการหาค่าความเชื่อมั่นของไฮเซนและบรอนสเตอร์ท (Heise and Bohmstedt. 1910) เมื่อพบว่าการกำหนดน้ำหนักความสำคัญไม่เท่ากันจะส่งผลกระทบต่อค่าความเชื่อมั่น และความคลาดเคลื่อน และได้เสนอสูตร Ω_w ซึ่งเป็นสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ เมื่อกำหนดน้ำหนักความสำคัญไม่เท่ากัน ไว้ดังนี้

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

- เมื่อ Ω_w แทน สัมประสิทธิ์ความเชื่อมั่น
 λ_i^2 แทน น้ำหนักองค์ประกอบของคำถามข้อที่ i

* จากสูตรการหาค่าความเชื่อมั่นดังกล่าว ในการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้สูตร r_{12} ของเดียว (บุญเชิด ภิญโญนันตพงษ์. 2537 : 33-37 ; อ้างอิงจาก Liou. 1989 : 153-163) และสูตร Ω_w ของอัลเลน (Allen. 1974) *

7. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวกับการศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธีในต่างประเทศ

เซนทรา(Centra. 1971 : 675-683) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธีของแบบทดสอบ IFI (Institutional Functioning Inventory) ที่วัดด้านสติปัญญา และการแสดงพฤติกรรมในชั้นเรียน ความเป็นอิสระ ความแตกต่างระหว่างบุคคล การเปลี่ยนแปลงทางสังคม การเรียนในระดับปริญญาตรี ความเป็นประชาธิปไตย ความสัมพันธ์กับบุคคลอื่น การรับรู้ตนเองและการวางแผน ความกระตือรือร้น ความเป็นคนทันสมัย ความมีปฏิภาณไหวพริบ โดยศึกษากับกลุ่มตัวอย่าง 3 กลุ่ม คือผู้บริหาร ครูผู้สอน และนิสิต กลุ่มละ 17 คน ผลการศึกษาพบว่า ในกลุ่มผู้บริหารมีความเชื่อมั่นตั้งแต่ .77 ถึง .84 กลุ่มครูผู้สอนมีความเชื่อมั่นตั้งแต่ .04 ถึง .84 และกลุ่มนิสิตมีความเชื่อมั่นตั้งแต่ .18 ถึง .85 ความเที่ยงตรงเชิงเหมือนมีค่าตั้งแต่ .20 ถึง .98 และมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

คาน (Khan. 1978 : 53 - 56) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี โดยใช้แบบประเมินเจตคติต่อโรงเรียนของนักเรียน ใน 4 ด้านคือ เจตคติต่อโรงเรียน เจตคติต่อครู เจตคติต่อตนเอง และเจตคติต่อสิ่งอื่น ๆ ใช้วิธีวัด 2 วิธี คือ ให้นักเรียนประเมินตนเอง และให้ครูเป็นผู้ประเมินเจตคติของนักเรียน โดยศึกษากับกลุ่มตัวอย่างที่เป็นนักเรียนชายหญิง อายุ 8 ปี จำนวน 387 คน ผลการศึกษาพบว่า ความเชื่อมั่นของแบบประเมินที่ให้นักเรียนประเมินตนเอง มีค่าตั้งแต่ .70 ถึง .84 ส่วนแบบประเมินโดยครูมีค่าตั้งแต่ .78 ถึง .93 ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

อาร์ลิน (Arlin. 1982 : 1077 - 1089) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี ของแบบทดสอบวัดความคิดของนักเรียน ใน 6 ด้านคือ การสนทนา การรวมกัน การจำแนกตัวแปร สัดส่วนความน่าจะเป็น และระบบการอ้างอิง ใช้วิธีวัด 2 วิธี คือ ให้นักเรียนเขียนในกระดาษคำตอบ และครูพูดคุยกับนักเรียนเป็นรายบุคคล กลุ่มตัวอย่างเป็นนักเรียนที่มีอายุเฉลี่ย 18 ปี จำนวน 38 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าตั้งแต่ .55 ถึง .74 ซึ่งสูงกว่าความเที่ยงตรงเชิงจำแนก สัมประสิทธิ์ค่าความเชื่อมั่นมีค่าตั้งแต่ .52 ถึง .79

มาร์ชและสมิท (Marsh and Smith. 1982 : 430-440) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธีของแบบทดสอบวัดความภูมิใจในตนเองของ คูเปอร์สมิท(SEI) เป็นแบบประเมินตนเอง 2 ระดับ คือ "ใช้ตัวฉัน" และ "ไม่ใช่ตัวฉัน" วัด 5 ด้านคือ ครอบครัว เพื่อน โรงเรียน กิจกรรมทั่วไป และความซื่อสัตย์ และแบบทดสอบการสำรวจภาพลักษณ์แห่งตนของเชียน (SCI) เป็นแบบมาตราส่วน 5 ระดับ วัด 9 ด้านคือ ความสามารถทางกาย ความพอใจในรูปร่างลักษณะ ความสัมพันธ์กับบุคคลอื่น ความรู้สึกที่ดีต่อสังคม ความสุขอุปนิสัยในการทำงาน กิจกรรมในโรงเรียน การคิดเอกลัษณ์ และการคิดนอกเนกนัย กลุ่มตัวอย่างเป็นนักเรียนเกรด 4,5 และ 6 ในออสเตรเลีย จำนวน 549 คน ผลการศึกษาพบว่า ความเชื่อมั่นของแบบทดสอบทั้ง 2 ฉบับมีค่าตั้งแต่ .63 ถึง .88 ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

ทริจันโควสกี(Tryjankowski. 1986 : 516) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี โดยใช้แบบประเมินระบบอาชีพของชาวยิว(JEVS) ใน 5 ด้านคือ การจำแนกทางการได้ยิน การจำแนกทางการเห็น การจดจำในสิ่งที่เห็น การจำภาพ การประสานกันของทักษะการเห็น และการประสานสัมพันธ์ของทักษะเครื่องมือวัด 2 ฉบับคือ แบบทดสอบวัดสติปัญญาของเวสเลอร์ฉบับปรับปรุง (WISC-R) และแบบทดสอบวัดความบกพร่องในการเรียน (SLDT) ซึ่งวัดลักษณะพิเศษของการรับรู้ กลุ่มตัวอย่างเป็นนักเรียนที่ลุ่มมาจาก

ความบกพร่องในการเรียน (SLDT) ซึ่งวัดลักษณะพิเศษของการรับรู้ กลุ่มตัวอย่างเป็นนักเรียนที่ส่งมาจาก ศูนย์ประเมินอาชีพในรัฐนิวเจอร์ซีย์ จำนวน 36 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

โมเซอร์(Crocker and Algina. 1986 : 233-234 ; citing Mosher. 1968 : 690-695) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี เมื่อกำหนดลักษณะที่วัด 3 ลักษณะ คือ ความรู้สึกทางเพศ ความรู้สึกผิดในการเป็นปรัักษ์ และความมีสติทางศีลธรรม ใช้วิธีวัด 3 วิธีคือ แบบทดสอบถูก-ผิด แบบเลือกตอบ และแบบเติมประโยคให้สมบูรณ์ กลุ่มตัวอย่างเป็นหญิงจำนวน 62 คน ผลการศึกษาพบว่า สัมประสิทธิ์สหสัมพันธ์ในรูปของความเชื่อมั่นมีค่าตั้งแต่ .41 ถึง .95 ความเที่ยงตรงเชิงเหมือนมีค่าตั้งแต่ .52 ถึง .90 และสูงกว่าความเที่ยงตรงเชิงจำแนก

แทรนท์ (Trant.1987:104) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี ของแบบทดสอบวัดลักษณะการเป็นผู้นำ (LST) ที่สร้างขึ้นเพื่อประเมิน ทักษะการเป็นผู้นำของนักเรียนระดับมัธยมศึกษา ใช้วิธีวัด 2 วิธีคือ ให้นักเรียนประเมินตนเอง และให้ครูเป็นผู้ประเมินนักเรียน กลุ่มตัวอย่างเป็นนักเรียนเกรด 11 และ 12 จำนวน 94 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

ซิงเกอร์(Singer.1989 : 275) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธีของแบบทดสอบวัดตนเอง(Self-Description Questionnaire : SDQ) วิธีวัด 3 วิธี คือให้นักเรียนประเมินตนเอง ผู้ปกครองประเมินนักเรียน และให้ครูเป็นผู้ประเมินนักเรียน กลุ่มตัวอย่างเป็นนักเรียนที่มีความบกพร่องทางร่างกาย จำนวน 169 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

เวนไรท์ (Wainwright. 1989 : 3006) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี ของแบบทดสอบวัดความสามารถทางสติปัญญาของเด็ก โดยใช้แบบทดสอบวัดสติปัญญาของเวสเลอร์ ฉบับปรับปรุง (WISC-R) และแบบทดสอบวัดความสามารถต่าง ๆ (DAS) และคะแนนจากแบบทดสอบวัดผลสัมฤทธิ์ของแอสตันฟอร์ด (SAT) นำมาจากระเบียนสะสมของนักเรียน กลุ่มตัวอย่างเป็นนักเรียนจากเขตชนบทหลายมาตตอนกลาง จำนวน 46 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

บาร์เบอร์ (Barbour. 1990 : 172) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธีของแบบประเมินพฤติกรรมของกลุ่มที่ไม่อยู่นิ่ง 3 ด้านคือ ความก้าวร้าว ความไม่สนใจในการเรียน และความวิตกกังวล ใช้วิธีประเมิน 3 วิธี คือ ให้นักเรียนประเมินตนเอง ผู้ปกครองประเมินนักเรียน และให้ครูเป็นผู้ประเมินนักเรียน กลุ่มตัวอย่างเป็นนักเรียนที่มีอายุระหว่าง 11-18 ปีจากศูนย์สุขภาพจิตชุมชน จำนวน 447 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

มาร์แชลล์ (Marshall. 1991 : 115) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี ของแบบทดสอบวัดกระบวนการเบื้องต้นทางวิทยาศาสตร์ ของนักเรียนระดับประถมศึกษา และระดับมัธยมศึกษา ใช้เครื่องมือวัด 4 ชนิด กลุ่มตัวอย่างเป็นนักเรียนเกรด 7 จำนวน 151 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

บัคเคิลว (Bucklew. 1988 : 106) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี ของแบบทดสอบวัดอาการทางประสาท 2 ฉบับคือ The Children's Depression Inventory(CDI) และ The Children's Action Tendency Scale (CATS) กลุ่มตัวอย่างเป็นนักเรียนที่อยู่หอพัก 4 แห่ง จำนวน 49 คนเป็นชาย 33 คน หญิง 16 คน ผลการศึกษาพบว่า ความเที่ยงตรงเชิงเหมือนมีค่าสูงกว่าความเที่ยงตรงเชิงจำแนก

งานวิจัยที่เกี่ยวกับการศึกษาคำความเที่ยงตรงแบบหลายลักษณะ-หลากวิธีในประเทศไทย

ตำเริง บุญเรืองรัตน์ (2528 : 95 - 100) ได้ศึกษาความเที่ยงตรงเชิงจำแนกของแบบทดสอบวัดลักษณะนิสัยการเป็นพยาบาล 5 ด้าน คือ ความมีเมตตา-กรุณา ความรับผิดชอบ ความขยันอดทน มนุษยสัมพันธ์ และการตัดสินใจ กลุ่มตัวอย่างเป็นนักเรียนที่สมัครเข้าศึกษาต่อที่วิทยาลัยพยาบาลเกื้อการุณย์ จำนวน 100 คน ผลการศึกษาพบว่า ค่าความเที่ยงตรงเชิงจำแนก ของแบบทดสอบวัดลักษณะนิสัยการเป็นพยาบาล 5 ด้าน มีค่าตั้งแต่ .06 ถึง .28

บรรทม มณีโชติ (2530 : 60-61) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธีของแบบทดสอบวัดลักษณะนิสัยด้านความเสียสละ 4 ด้าน คือ การให้ปัน การเห็นแก่ส่วนรวม ความมีน้ำใจ และการไม่เอาเปรียบผู้อื่น ใช้วิธีวัด 2 วิธี คือแบบทดสอบแบบข้อความ และแบบสถานการณ์ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2539 ของโรงเรียนสังกัดสำนักงานการประถมศึกษา จังหวัดเพชรบุรี จำนวน 83 คน ผลการศึกษาพบว่า ค่าความเที่ยงตรงเชิงเหมือน มีค่าตั้งแต่ .4116 ถึง .6104 มีนัยสำคัญทางสถิติทุกค่า และความเที่ยงตรงเชิงจำแนก มีค่าตั้งแต่ -.0270 ถึง .2155

สุภาวดี ตั้งบุปผา (2533 : 101-102) ได้ศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี ของแบบทดสอบวัดความสามารถในการคิดสร้างสรรค์ทางคณิตศาสตร์ 7 ด้าน ใช้วิธีวัด 2 วิธี คือ แบบทดสอบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์ที่ผู้วิจัยสร้างเอง และแบบทดสอบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์ของทอร์เรนซ์ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดกรมสามัญศึกษา เขตการศึกษา 2 กรุงเทพมหานคร จำนวน 367 คน ผลการศึกษาพบว่า ค่าความเที่ยงตรงเชิงเหมือน มีค่าตั้งแต่ .7089 ถึง .8794 มีนัยสำคัญทางสถิติทุกค่า และความเที่ยงตรงเชิงจำแนก มีค่าตั้งแต่ .1538 ถึง .3015

พรรณี เทพสุทร (2537 : 68) ได้ศึกษาความเที่ยงตรงเชิงโครงสร้างของแบบวัดความวิตกกังวลในการเรียนวิชาคณิตศาสตร์ 4 ด้าน ใช้วิธีวัด 2 วิธี คือ แบบทดสอบชนิดข้อความ และแบบทดสอบชนิดสถานการณ์ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 สังกัดกรมสามัญศึกษา ในจังหวัดเพชรบุรี จำนวน 536 คน ผลการศึกษาพบว่า ค่าความเที่ยงตรงเชิงเหมือน มีค่าตั้งแต่ .5961 ถึง .7769 และความเที่ยงตรงเชิงจำแนก มีค่าตั้งแต่ -.3602 ถึง .4218 (เขียน 4 มิถุนายน 2537)

เบญจวรรณ แยมละมูล (2538 : 61 - 62) ได้ศึกษาความเที่ยงตรงเชิงโครงสร้างของแบบวัดเจตคติที่มีต่อโรงเรียนของนักเรียนและครูในระดับประถมศึกษา กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 สังกัดสำนักงานการประถมศึกษา จังหวัดนครปฐม จำนวน 769 คน และกลุ่มตัวอย่างที่เป็นครูสังกัดโรงเรียนประถมศึกษา จังหวัดนครปฐม จำนวน 390 คน เครื่องมือที่ใช้ เป็นแบบทดสอบวัดเจตคติที่มีต่อโรงเรียนในด้านต่าง ๆ 4 ด้าน แบบลิเคิร์ท และแบบออสกูด ผลการศึกษาพบว่า ค่าความเที่ยงตรงเชิงเหมือนของกลุ่มนักเรียน มีค่าตั้งแต่ .7823 ถึง .9378 และ การวิเคราะห์จากกลุ่มครูมีค่าตั้งแต่ .7091 ถึง .9151 ส่วนค่าความเที่ยงตรงเชิงจำแนกของกลุ่มนักเรียน มีค่าตั้งแต่ .306 ถึง .4543 และการวิเคราะห์จากกลุ่มครูมีค่าตั้งแต่ -.3325 ถึง .4783

เสกสิทธิ์ แสนทวีสุข (2539 : 80 - 81) ได้ศึกษาความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบวัดมโนภาพแห่งตนหลายมิติ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2538 ของโรงเรียนสังกัดคณะกรรมการการประถมศึกษาแห่งชาติ ที่เป็นโรงเรียนขยายโอกาสทางการศึกษา จังหวัดศรีสะเกษ จำนวน 514 คน เครื่องมือที่ใช้เป็นแบบทดสอบที่แปลมาจากแบบทดสอบของมาร์ช และแบบทดสอบที่ผู้วิจัยสร้างขึ้น ประกอบด้วย 11 ด้าน ผลการ

ความเที่ยงตรงเชิงเหมือน มีค่าตั้งแต่ .5211 ถึง .7982 ซึ่งมีค่าสูงกว่า .50 ตามต้องการ และค่าความเที่ยงตรงเชิงจำแนกมีค่าตั้งแต่ -.2179 ถึง .4762 ซึ่งต่ำกว่าความเที่ยงตรงเชิงเหมือนตามที่ระบุไว้

งานวิจัยที่เกี่ยวกับข้อคำถามปลายเปิด

กรอนลันด์ (Gronlund. 1965 : 108) กล่าวว่าแบบทดสอบปรนัยใช้วัดความรู้เกี่ยวกับความจริง ความเข้าใจ ทักษะในการคิด เหมาะในการวัดความจำ แต่แบบทดสอบอัตนัยใช้ได้ในการวัดความเข้าใจ ทักษะ การคิด ความสามารถในการเขียน การแก้ปัญหา โดยเฉพาะวัดความคิดริเริ่มสร้างสรรค์ และเหมาะสมอย่างยิ่ง กับคำถามที่ให้นักเรียนใช้ความสามารถในการหาคำตอบเอง

ไทเลอร์ (Cronbach. 1960 : 373 ciling Tyler. 1934) ได้ทำการศึกษากับนักศึกษาระดับต่ำกว่า ปริญญาตรีในด้านการสรุปความทั่วไป โดยใช้แบบทดสอบปรนัย และแบบทดสอบอัตนัย ผลปรากฏว่า ค่า สหสัมพันธ์ของคะแนนจากแบบทดสอบเลือกตอบ กับแบบทดสอบอัตนัย เท่ากับ .85 และมีค่าเท่ากับ .79 จากการ สอบวัดในวิชาการวางแผนงาน

ไบเรเนบอม และทัซสุโอกะ (Birenbaum and Tasuoka. 1987 : 385-395) ได้ทำการวิจัยเพื่อ ตรวจสอบผลของการใช้รูปแบบการตอบสนองในข้อคำถามที่เป็นแบบปลายเปิด (Open-Ended) กับแบบ เลือกตอบ (Multiple-Choice) ว่าเมื่อนำไปใช้ในการ วิจัยความเข้าใจผิดในการงานที่มีลักษณะเป็นวิธีการ แล้วข้อมูล 2 ชุด ที่ได้จากข้อคำถาม 2 รูปแบบ สามารถสะท้อนให้เห็นจำนวนชนิดของความคลาดเคลื่อนและ วินิจฉัยแหล่งของความเข้าใจผิดได้เพียงใด กลุ่มตัวอย่างเป็นนักเรียนเกรด 8 จำนวน 285 คน ทำข้อสอบ ปลายเปิด 148 คน และทำข้อสอบแบบเลือกตอบ จำนวน 137 คน นักเรียนทั้ง 2 กลุ่มเป็นนักเรียนที่ เรียนปฏิบัติการทดลองทางคณิตศาสตร์เป็นเวลาติดต่อกัน 2 ปี แบบทดสอบปลายเปิด ประกอบด้วยข้อคำถาม 38 ข้อ จัดเป็นแบบทดสอบคู่ขนาน ชุดละ 19 ข้อ แบบทดสอบเลือกตอบ มี 5 ตัวเลือก ตัวลวงสร้างจาก คำตอบของนักเรียนที่ตอบผิดจากแบบทดสอบปลายเปิดที่ศึกษาในลักษณะนำร่องไปแล้ว ผลการศึกษาแสดงให้เห็นว่า แม้กลุ่มผู้สอบทั้งสองกลุ่มจะไม่แตกต่างกันในความสามารถในการแก้ปัญหาการบวกเศษส่วน แต่ข้อ คำถามปลายเปิดก็ให้ผลเกี่ยวกับ การพิจารณาจำนวนชนิดของความคลาดเคลื่อน และวินิจฉัยแหล่งของความ เข้าใจผิดได้เป็นที่น่าพอใจมากกว่า

มาเรียและคณะ (Maria E. Magone, Jinfa Cai, Edward A. Silver, and Ning Wang. 1994 : 339) ได้ศึกษาการประเมินการแสดงวิธีทำ การให้เหตุผลในความคิดซับซ้อน และขอบเขตเนื้อหาวิชาคณิตศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนเกรด 6 และเกรด 7 จำนวน 1,627 คน แบ่งเป็น 3 กลุ่ม คือ

กลุ่มที่ 1 สอบเกี่ยวกับความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ จำนวน 568 คน

กลุ่มที่ 2 สอบเกี่ยวกับการแก้โจทย์ปัญหาทางคณิตศาสตร์ จำนวน 485 คน และ

กลุ่มที่ 3 สอบเกี่ยวกับการให้เหตุผลทางคณิตศาสตร์ จำนวน 574 คน

เครื่องมือที่ใช้เป็นแบบทดสอบปลายเปิดจำนวน 9 ข้อ ประกอบด้วยการประเมิน

1. ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์

2. ความสามารถในการแก้โจทย์ปัญหา

3. การให้เหตุผลทางคณิตศาสตร์

ผลการวิเคราะห์คำตอบของนักเรียนพบว่า

1. งานสอบที่ประเมินความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ พบว่า นักเรียนตอบถูก 45% และนักเรียนเกรด 6 จะตอบถูกเพิ่มขึ้นเป็น 59% เมื่ออยู่เกรด 7 ครึ่งหนึ่งของนักเรียนที่ตอบถูกให้ เหตุผลได้ถูกต้องหรือผิดพลาดเล็กน้อย นักเรียนที่ตอบผิดส่วนใหญ่จะเลือกตอบ 0.008000 คิดเป็น 37 %

ของนักเรียนทั้งหมด เป็นนักเรียนเกรด 6 ภาคเรียนที่ 2 50 % และเกรด 7 ภาคเรียนที่ 1 27 % และส่วนใหญ่ใช้การเขียนบรรยาย และให้เหตุผลในรูปของความสัมพันธ์ระหว่างตัวเลขและตำแหน่งของทศนิยมหรือจำนวนเต็ม หรือใช้เศษส่วนที่เทียบเท่ากัน

2. งานสอบที่ประเมินความสามารถในการแก้โจทย์ปัญหา พบว่า ครึ่งหนึ่งนักเรียนใช้รูปภาพแสดงการแก้ปัญหา นอกนั้นใช้ประโยคสัญลักษณ์ทางคณิตศาสตร์ การบรรยาย และใช้ร่วมกันระหว่างการบรรยาย รูปภาพและสัญลักษณ์ นักเรียนใช้กลยุทธ์ในการแก้ปัญหาที่หลากหลายไม่น้อยกว่า 5 วิธี ข้อบกพร่องจากคำตอบของนักเรียนที่พบ คือ ความไม่เข้าใจในเงื่อนไขของปัญหา

3. งานสอบที่ประเมินการให้เหตุผลทางคณิตศาสตร์ พบว่า นักเรียนส่วนใหญ่เขียนรูปต่อไปได้และให้เหตุผลด้วยวิธีบรรยาย 90 % นอกนั้น ใช้รูปภาพ หรือจุดในการแสดงคำตอบ ข้อบกพร่องที่พบ คือ 40 % ที่เขียนรูปไม่ถูกต้อง เนื่องจากความผิดพลาดในการใช้ความสัมพันธ์ของภาพ 2 มิติ จำนวนจุดและรูปร่างของภาพ 30 % ของกลุ่มที่เขียนรูปไม่ถูกต้องเนื่องจาก รูปร่างของภาพไม่ถูก แต่จำนวนจุดถูก หรือจำนวนจุดไม่ถูกแต่รูปร่างถูกต้อง และ 54 % ของกลุ่มที่รูปไม่ถูกต้องที่ผิดทั้งรูปร่างและจำนวนจุด

จากการศึกษาทั้งหมดได้ข้อสรุปดังนี้

1. นักเรียนใช้ตัวอย่างที่หลากหลาย
2. นักเรียนแสดงวิธีแก้ปัญหาได้หลายรูปแบบ
3. ความเข้าใจผิดและข้อบกพร่องพบได้จากการให้เหตุผลและขั้นตอนการแก้ปัญหา
4. คำตอบของนักเรียนแสดงให้เห็นถึงความลึกในขอบเขตของเนื้อหาวิชา
5. ในขั้นที่สูงขึ้นคำตอบของนักเรียนจะสมบูรณ์มากขึ้น

เลน และคณะ (Lane and Others, 1994 : 247 - 261) ได้ศึกษาเกี่ยวกับค่าความเชื่อมั่นและความเที่ยงตรงของการประเมิน วิธีทำวิชาคณิตศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนเกรด 6 และเกรด 7 จาก 6 โรงเรียนที่เข้าร่วมโครงการ QUASAR ของประเทศสหรัฐอเมริกา จำนวน 1,579 คนเครื่องมือที่ใช้เป็นแบบทดสอบปลายเปิด 4 ฉบับๆละ 9 ข้อ แต่ละฉบับจะประเมินการให้เหตุผลทางคณิตศาสตร์ การแก้โจทย์ปัญหา ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ การสื่อสารทางคณิตศาสตร์ พบว่า สัมประสิทธิ์การสรุปอ้างอิง เมื่อ $n = 9$ มีค่าตั้งแต่ .71 ถึง .84

สมสว่าง วัฒพานิชย์สกุล (2539 : 49-51) ได้สร้างแบบทดสอบวัดกระบวนการแก้ปัญหาโจทย์คณิตศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 โรงเรียนกุนนที่รุตธารามวิทยาคม สังกัดกรมสามัญศึกษา กรุงเทพมหานคร ที่ลงทะเบียนเรียนวิชา ค 015 จำนวน 250 คน ผลการศึกษาพบว่า ค่าความยากของข้อสอบมีค่า .32 ถึง .62 ค่าอำนาจจำแนกของข้อสอบมีค่า .34 ถึง .79 และค่าความเชื่อมั่นของแบบทดสอบมีค่า .8068

ปราโมทย์ มากชู (2540 : 28-39) ได้ศึกษาพฤติกรรมการคิดแก้โจทย์ปัญหาคณิตศาสตร์ของนักเรียน ระดับชั้นมัธยมศึกษาปีที่ 3 ในเขตภาคเหนือตอนล่าง กลุ่มตัวอย่างเป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 จากโรงเรียนสังกัดกรมสามัญศึกษาในเขตภาคเหนือตอนล่าง จำนวน 400 คน เครื่องมือที่ใช้เป็นแบบทดสอบอัตนัย และการสัมภาษณ์พร้อมบันทึกเทปวีดิทัศน์ ผลการวิจัยพบว่า นักเรียนใช้พฤติกรรมการคิดแก้โจทย์ปัญหาคณิตศาสตร์ ตามบทนิยาม ตามสมบัติ ตามบทนิยามและสมบัติ และตามแบบอิสระ คิดเป็นร้อยละ 14.95, 47.31, 37.37 และ .37 ตามลำดับ

จากงานวิจัยดังกล่าว จะเห็นว่าข้อคำถามปลายเปิดสามารถวัดกระบวนการคิดที่ซับซ้อนได้ดี เหมาะกับแบบทดสอบที่ต้องการวัดกระบวนการคิด ซึ่งนักเรียนสามารถแสดงการคิดได้อย่างอิสระ ได้รูปแบบการตอบที่หลากหลาย และสามารถวัดความคิดของนักเรียนได้โดยตรง ผู้วิจัยจึงสนใจที่จะศึกษาคุณภาพของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ว่ามีความเที่ยงตรงเชิงโครงสร้าง และความเชื่อมั่นเป็นอย่างไร

สมมติฐานของการศึกษาค้นคว้า

1. ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่ใช้วิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) จะมีความเที่ยงตรงเชิงเหมือนสูงกว่าความเที่ยงตรงเชิงจำแนก
2. ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของไฮเซน และบรอนสเตอร์ (Heise and Bohmstedt, 1970) มีค่าสูงกว่า .5
3. ความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร Ω_w ของอัลเลน (Allen, 1974) มีค่าสูงกว่า .5
4. ความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร r_{L2} ของเลียว (Liou, 1989) มีค่าสูงกว่า .5

บทที่ 3

วิธีดำเนินการศึกษาค้นคว้า

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 สังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพมหานคร ที่เป็นโรงเรียนสหศึกษาจำนวน 11 โรงเรียน 159 ห้องเรียน จำนวนนักเรียน 6,987 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 สังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพมหานคร ที่เป็นโรงเรียนสหศึกษา ซึ่งเลือกมาโดยการสุ่มแบบ 2 ขั้นตอน(Two-Stage Cluster Sampling) ดังนี้

ขั้นที่ 1 สุ่มอย่างง่าย (Simple Random Sampling) โดยใช้โรงเรียนเป็นหน่วยการสุ่ม จำนวน 3 โรงเรียน

ขั้นที่ 2 สุ่มอย่างง่าย (Simple Random Sampling) อีกครั้ง โดยใช้ห้องเรียนเป็นหน่วยการสุ่ม จำนวน 6 ห้องเรียน ได้จำนวนนักเรียน 300 คน ดังแสดงรายละเอียดในตาราง 3

ตาราง 3 จำนวนห้องเรียนและจำนวนนักเรียนที่เป็นกลุ่มตัวอย่างจำแนกตามโรงเรียน

โรงเรียน	จำนวนนักเรียนในกลุ่มตัวอย่าง	
	ห้องเรียน	จำนวนนักเรียน
โรงเรียนพระโขนงพิทยาลัย	2	94
โรงเรียนสุรศักดิ์มนตรี	2	95
โรงเรียนเตรียมอุดมศึกษาพัฒนาการ	2	111
รวม	6	300

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้เป็นแบบทดสอบ 2 ฉบับ คือ

1. แบบทดสอบปลายเปิด
2. แบบทดสอบเลือกตอบ

เนื้อหาที่ใช้ออกข้อสอบเป็นเนื้อหาวิชาคณิตศาสตร์ (ค203) ชั้นมัธยมศึกษาปีที่ 2 จากหนังสือแบบเรียนวิชาคณิตศาสตร์ ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) เรื่องระบบจำนวนเต็ม เศษส่วนและทศนิยม

ลักษณะของเครื่องมือ

เครื่องมือที่ใช้ในการศึกษาครั้งนี้เป็นแบบทดสอบ 2 ฉบับ ดังนี้

1. แบบทดสอบปลายเปิด วัดพฤติกรรมการแสดงออกถึงความคิดซับซ้อน จำนวน 10 ข้อ แบ่งเป็น 3 ด้าน ดังนี้

- | | |
|--|-------|
| 1.1 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ | 3 ข้อ |
| 1.2 การแก้ปัญหา | 4 ข้อ |
| 1.3 การให้เหตุผล | 3 ข้อ |

2. แบบทดสอบเลือกตอบ ชนิด 5 ตัวเลือก สร้างจากข้อคำถามปลายเปิดของแบบทดสอบปลายเปิด จำนวน 30 ข้อ แบ่งเป็น 3 ด้าน ดังนี้

- | | |
|--|--------|
| 2.1 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ | 10 ข้อ |
| 2.2 การแก้ปัญหา | 10 ข้อ |
| 2.3 การให้เหตุผล | 10 ข้อ |

ตัวอย่างแบบทดสอบปลายเปิด

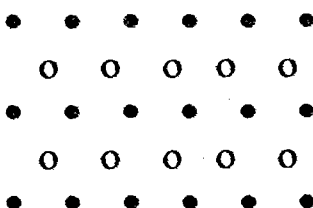
คำชี้แจง 1. แบบทดสอบฉบับนี้ประกอบด้วยข้อคำถามปลายเปิด จำนวน 10 ข้อ

2. ในแต่ละข้อให้นักเรียนแสดงวิธีคิดพร้อมทั้งแสดงเหตุผลประกอบการคิดในรูปแบบต่าง ๆ เช่น การบรรยาย การเขียนแผนภูมิ รูปภาพ กราฟ ตาราง หรือสัญลักษณ์ทางคณิตศาสตร์ เป็นต้น

3. ในแต่ละข้อให้นักเรียนเขียนตอบในกระดาษคำตอบที่กำหนดให้

ข้อ(0)

ลูกแก้วสีดำและสีขาวถูกจัดเรียงดังรูป



จากรูปข้างบน จงหาว่ามีลูกแก้วทั้งหมดเท่าใดและใช้หลักการใดในการหาคำตอบ เพราะเหตุใด(บอกมา 2 วิธี)

เกณฑ์การให้คะแนน

- 0 คะแนน ไม่ตอบหรือตอบไม่ถูกเลย
- 1 คะแนน บอกหลักการที่ใช้ในการหาจำนวนลูกแก้วทั้งหมดได้เล็กน้อย แต่ไม่ได้คำตอบ หรือคำตอบไม่ถูกต้อง
- 2 คะแนน บอกหลักการ และหาคำตอบตามหลักการได้ถูกต้อง 1 วิธี
- 3 คะแนน คำตอบถูกต้อง อธิบายหลักการได้เกือบสมบูรณ์
- 4 คะแนน คำตอบถูกต้อง อธิบายหลักการได้ถูกต้องสมบูรณ์ทั้ง 2 วิธี

ข้อ (00)

“สามีของอารยา เขียนพินัยกรรมไว้ว่า ถ้าภรรยาของเขาตลอดลูกชาย ให้แบ่งมรดกให้ลูกชาย $\frac{2}{3}$ ของทั้งหมด ที่เหลือให้ภรรยา ถ้าตลอดลูกสาวให้แบ่งมรดกให้ลูกสาว $\frac{1}{5}$ ของทั้งหมด ที่เหลือให้ภรรยา”

ต่อมาปรากฏว่า ภรรยาของเขาตลอดลูกแฝดเป็นชาย 2 คน หญิง 1 คน จงหาว่า ภรรยา และลูกแต่ละคน จะได้รับมรดกตามเงื่อนไขดังกล่าวคนละเท่าใดของทั้งหมด

.....

.....

.....

.....

.....

ข้อ (000)

จากรูปแบบที่กำหนดให้ ใช้ตอบคำถามข้อ 1-2

จำนวนของตัวเลข	เขียนในรูปการบวก	ผลลัพธ์
1	1	1
2	1+3	4
3	1+3+5	9
4	1+3+5+7	16
---	-----	-----

1. จำนวนในแถวถัดไปคือ

.....

2. ความสัมพันธ์ระหว่างจำนวนของตัวเลขในคอลัมน์ที่ 1 และผลรวมในคอลัมน์ที่ 2 คือ

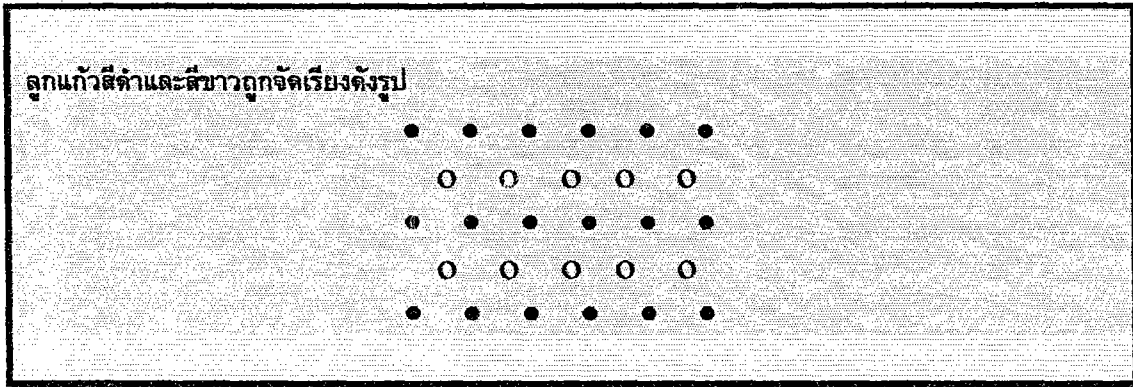
.....

.....

.....

ตัวอย่าง แบบทดสอบเลือกตอบ

- คำชี้แจง 1. แบบทดสอบมีทั้งหมด 30 ข้อ คะแนนเต็ม 30 คะแนน
 2. แต่ละข้อมีคำตอบถูกเพียงคำตอบเดียว
 3. ให้นักเรียนทำเครื่องหมาย X ทับตัวอักษร ก ,ข ,ค ,ง และ จ ที่ถูกต้องที่สุดเพียงคำตอบเดียว
 ในกระดาษคำตอบที่กำหนดให้



จากรูปใช้ตอบคำถามข้อ 1-3

1. นักเรียนจะหาจำนวนลูกแก้วทั้งหมดโดยวิธีใด
 - ก. a^n
 - ข. การคูณ
 - ค. การเพิ่มครั้งละ 6
 - ง. a^n และการบวก
 - จ. การคูณ และการบวก
2. ข้อใดถูกต้อง
 - ก. มีลูกแก้ว ● เป็น 2 เท่าของลูกแก้ว ○
 - ข. มีลูกแก้ว ○ เป็น 2 เท่าของลูกแก้ว ●
 - ค. มีลูกแก้ว ● มากกว่าลูกแก้ว ○ อยู่ 2^3
 - ง. มีลูกแก้ว ○ อยู่ 2 ใน 5 ของลูกแก้วทั้งหมด
 - จ. ถ้าเพิ่มลูกแก้วที่ละคู่ จะมีลูกแก้วทั้งหมดเป็นจำนวนคี่เสมอ
3. จากรูป มีลูกแก้วทั้งหมด 28 ลูก เพราะเหตุใด
 - ก. 2×14
 - ข. $2(8 + 6)$
 - ค. $4(4 + 3)$
 - ง. $(6 \times 3) + (5 \times 2)$
 - จ. $(5 \times 4) + (2 \times 4)$

“สามีของอารยา เขียนพินัยกรรมไว้ว่า “ถ้าภรรยาของเขาคลอลูกชาย ให้แบ่งมรดกให้ลูกชาย $\frac{2}{3}$ ของทั้งหมดที่เหลือให้ภรรยา แต่ถ้าคลอลูกสาว ให้แบ่งมรดกให้ลูกสาว $\frac{1}{5}$ ของทั้งหมด ที่เหลือให้ภรรยา”

ต่อมาปรากฏว่า ภรรยาของเขาคลอลูกแฝดเป็นชาย 2 คน หญิง 1 คน จงหาว่า ภรรยา และลูกแต่ละคนจะได้รับมรดกคนละเท่าใด

จากโจทย์ใช้ตอบคำถามข้อ 4-6

4. ภรรยาจะได้รับมรดกเท่าใด หลังจากแบ่งให้ลูกสาว และลูกชายแต่ละคนแล้ว

- ก. $1 - \frac{2}{3}$ ของทั้งหมด
- ข. $1 - (\frac{2}{3} + \frac{1}{5})$ ของทั้งหมด
- ค. $1 - (\frac{2}{3} + \frac{2}{3})$ ของทั้งหมด
- ง. $1 - (\frac{2}{3} - \frac{1}{5})$ ของทั้งหมด
- จ. $1 - (\frac{2}{3} + \frac{2}{3} + \frac{1}{5})$ ของทั้งหมด

5. ภรรยา ลูกสาว และลูกชายแต่ละคน จะได้รับมรดกคนละเท่าใด

- ก. $\frac{1}{3}, \frac{1}{5}, \frac{2}{3}, \frac{2}{3}$ ตามลำดับ
- ข. $\frac{1}{3}, \frac{1}{5}, \frac{1}{3}, \frac{1}{3}$ ตามลำดับ
- ค. $\frac{4}{5}, \frac{1}{5}, \frac{1}{3}, \frac{1}{3}$ ตามลำดับ
- ง. $\frac{4}{5}, \frac{1}{5}, \frac{2}{3}, \frac{2}{3}$ ตามลำดับ
- จ. $\frac{2}{15}, \frac{1}{5}, \frac{1}{3}, \frac{1}{3}$ ตามลำดับ

6. ลูกสาวและลูกชายจะได้รับมรดกเท่ากันเมื่อใด

- ก. แม่ยกมรดกของตนเองทั้งหมดให้ลูกสาว
- ข. แม่แบ่งมรดกให้ลูกสาวครึ่งหนึ่งของที่มีอยู่
- ค. แม่แบ่งมรดกของตนเองทั้งหมดให้กับลูกแต่ละคน ๆ ละเท่าๆกัน
- ง. แม่และลูกสาว แบ่งมรดกให้ลูกชายแต่ละคน ๆ ละ หนึ่งในสี่ ของที่มีอยู่
- จ. แม่และลูกชายแต่ละคน แบ่งมรดกให้ลูกสาว คนละ หนึ่งในสี่ ของที่มีอยู่

จำนวนของตัวเลข	เขียนในรูปผลบวก	ผลลัพธ์
1	1	1
2	1+3	4
3	1+3+5	9
4	1+3+5+7	16
.....		
.....		
.....		

จากรูปแบบที่กำหนดให้ ใช้ตอบคำถามข้อ 7-9

7. จากรูปแบบที่กำหนดให้ ข้อใดถูกต้อง

- ก. เป็นการหาผลรวมของ จำนวนนับ
- ข. เป็นการหาผลรวมของ จำนวนเต็ม
- ค. เป็นการหาผลรวมของ จำนวนคี่บวก
- ง. เป็นการหาผลรวมของ จำนวนเฉพาะ
- จ. เป็นการหาผลรวมของ ตัวประกอบเฉพาะ

8. จากรูปแบบ จงหาผลรวมในแถวที่ 10

- ก. 60
- ข. 72
- ค. 88
- ง. 100
- จ. 121

9. ถ้า n แทนตัวเลขในคอลัมน์ที่ 1 จงหาความสัมพันธ์ระหว่างคอลัมน์ที่ 1 และคอลัมน์ที่ 3

- ก. 2^n
- ข. n^2
- ค. $2n$
- ง. $n + 2$
- จ. $2n + 1$

วิธีดำเนินการสร้างเครื่องมือ

การสร้างเครื่องมือ เพื่อใช้ในการเก็บรวบรวมข้อมูล ผู้วิจัยได้ดำเนินการสร้างแบบทดสอบ ตามลำดับขั้นตอนดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ ดังนี้

1.1 สร้างแบบทดสอบวิชาคณิตศาสตร์แบบปลายเปิด และแบบเลือกตอบ โดยใช้เนื้อหาวิชาคณิตศาสตร์ (ค203) ชั้นมัธยมศึกษาปีที่ 2 จากหนังสือแบบเรียนวิชาคณิตศาสตร์ ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) โดยสร้างแบบทดสอบเลือกตอบ จากข้อคำถามของแบบทดสอบปลายเปิด

1.2 เพื่อศึกษาคุณภาพของแบบทดสอบ

2. ศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับคำถามปลายเปิด การสร้างข้อคำถามปลายเปิด เกณฑ์การให้คะแนน และการสร้างแบบทดสอบเลือกตอบ เพื่อเป็นแนวทางในการสร้างแบบทดสอบ(ดังปรากฏในหนังสืออ้างอิง)

3. ศึกษาหลักสูตร คู่มือครู หนังสือแบบเรียนวิชาคณิตศาสตร์ (ค203) ชั้นมัธยมศึกษาปีที่ 2 จากหนังสือแบบเรียนวิชาคณิตศาสตร์ ตามหลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533) เพื่อกำหนดขอบเขตของเนื้อหา พฤติกรรมที่ต้องการวัด

4. นิยามปริเขตโครงสร้างที่ต้องการวัด ในที่นี้คือ ขอบเขตของเนื้อหาที่ใช้ประเมินพฤติกรรมที่แสดงออกถึงการคิดซับซ้อน เกี่ยวกับความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ การแก้ปัญหา การให้เหตุผล

5. วิเคราะห์ปริเขตโครงสร้างที่ต้องการวัด เพื่อกำหนดลักษณะเฉพาะ และกระบวนการ(พฤติกรรม) ที่ต้องการประเมิน ดังแสดงรายละเอียดในตาราง 4

ตาราง 4 ลักษณะเฉพาะของแบบทดสอบ

ส่วนประกอบหลัก	ส่วนประกอบย่อย
1. เนื้อหา	- ระบบจำนวนเต็ม - เศษส่วนและทศนิยม
2. กระบวนการทางความคิด	- ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ - ความสามารถในการแก้ปัญหาทางคณิตศาสตร์ - ความสามารถในการให้เหตุผลทางคณิตศาสตร์
3. รูปแบบการนำเสนอแนวคิด	- คำอธิบาย - เขียนแผนผังหรือภาพประกอบ - สร้างตารางหรือกราฟ - การเดาและการตรวจสอบ - การแจกแจงกรณีที่เป็นไปได้ทั้งหมด - การใช้สัญลักษณ์หรือประโยคทางคณิตศาสตร์ - การหารูปแบบ

6. เขียนข้อคำถามปลายเปิดตามนิยามปริเขตโครงสร้าง จำนวน 17 ข้อ
7. นำแบบทดสอบที่ได้ในข้อ 6 ไปทดสอบครั้งที่ 1 กับกลุ่มตัวอย่างจำนวน 30 คน เพื่อสำรวจข้อบกพร่องและสร้างเกณฑ์การให้คะแนน
8. นำแบบทดสอบที่ปรับปรุงจากข้อ 7 ให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงตามเนื้อหา และเกณฑ์การให้คะแนน
9. ปรับปรุงแก้ไขข้อสอบ โดยอาศัยข้อมูลที่ได้จากข้อ 8 แล้วจัดทำเป็นแบบทดสอบนำไปสอบครั้งที่ 2 กับกลุ่มตัวอย่างจำนวน 50 คน เพื่อหาค่าความยากและค่าอำนาจจำแนก แล้วคัดเลือกข้อที่มีค่าความยากตั้งแต่ .46 ถึง .62 และค่าอำนาจจำแนกที่มีนัยสำคัญทางสถิติ (โดยทดสอบค่าสหสัมพันธ์แบบเพียร์สัน) มีค่าตั้งแต่ .31 ถึง .59 ไว้จำนวน 10 ข้อ
10. จัดพิมพ์ข้อสอบที่ผ่านการคัดเลือกในข้อ 9 พร้อมทั้งเกณฑ์การให้คะแนน
11. นำข้อสอบที่ได้จากข้อ 10 ไปออกข้อสอบแบบเลือกตอบ ชนิด 5 ตัวเลือก จำนวน 45 ข้อนำไปให้ผู้เชี่ยวชาญตรวจสอบความเที่ยงตรงตามเนื้อหา
12. ปรับปรุงแก้ไขข้อสอบ โดยอาศัยข้อมูลที่ได้จากข้อ 11 แล้วจัดทำเป็นแบบทดสอบ เพื่อนำไปสอบกับกลุ่มตัวอย่างกลุ่มเดียวกันกับที่สอบแบบทดสอบปลายเปิดครั้งที่ 2 จำนวน 50 คน เพื่อหาค่าความยากและค่าอำนาจจำแนก โดยใช้โปรแกรม Testqual แล้วคัดเลือกข้อที่มีค่าความยากตั้งแต่ .26 ถึง .72 และค่าอำนาจจำแนกตั้งแต่ .20 ถึง .63 ไว้จำนวน 30 ข้อ
13. จัดพิมพ์ข้อสอบที่ผ่านการคัดเลือกในข้อ 12
14. นำแบบทดสอบที่ได้ในข้อ 10 และข้อ 13 ไปสอบกับกลุ่มตัวอย่าง จำนวน 300 คน เพื่อแสดงค่าความเที่ยงตรงเชิงโครงสร้าง และค่าความเชื่อมั่นของแบบทดสอบ

วิธีการเก็บรวบรวมข้อมูล

การดำเนินการเก็บรวบรวมข้อมูล ผู้วิจัยมีวิธีดำเนินการดังต่อไปนี้

1. ทำหนังสือขอความร่วมมือจากผู้อำนวยการโรงเรียนที่เป็นกลุ่มตัวอย่าง ในการเก็บรวบรวมข้อมูล เพื่อกำหนดวัน เวลา และสถานที่สอบ
2. ชี้แจงให้นักเรียนกลุ่มตัวอย่าง ทราบวัตถุประสงค์ของการสอบ และขอความร่วมมือในการสอบ เพื่อให้ได้ผลตามความเป็นจริง
3. ทำการทดสอบกับนักเรียนกลุ่มตัวอย่าง ด้วยข้อสอบทั้ง 2 ฉบับ โดยให้นักเรียนสอบด้วยแบบทดสอบปลายเปิดก่อน หลังจากนั้นประมาณ 1 สัปดาห์จึงสอบด้วยแบบทดสอบเลือกตอบ
4. ตรวจให้คะแนนตามเกณฑ์ที่กำหนด
5. นำผลที่ได้มาทดสอบค่าสถิติ

ขั้นตอนการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลมีขั้นตอน ดังนี้

1. ตรวจให้คะแนนตามเกณฑ์ที่กำหนด
2. หาค่าสถิติพื้นฐาน ของคะแนนจากการสอบทั้ง 2 ฉบับ
3. หาคุณภาพของแบบทดสอบ

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลใช้วิธีการหาค่าสถิติ ดังนี้

1. หาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ยเลขคณิต และค่าความเบี่ยงเบนมาตรฐาน
2. ค่าความยากและค่าอำนาจจำแนกรายข้อของแบบทดสอบปลายเปิดหาจาก
 - 2.1 ค่าความยากรายข้อ(บุญชม ศรีสะอาด. 2532)

$$p_i = \frac{\bar{x}_i}{k_i}$$

เมื่อ	p_i	แทน	ค่าความยากของข้อสอบข้อที่ i
	x_i	แทน	คะแนนเฉลี่ยของข้อสอบข้อที่ i
	k_i	แทน	คะแนนเต็มของข้อสอบข้อที่ i

2.2 ค่าอำนาจจำแนกรายข้อของแบบทดสอบ จากค่าสหสัมพันธ์ระหว่างคะแนนของข้อสอบแต่ละข้อ กับคะแนนรวม ดำเนินโดยวิธีสัมประสิทธิ์แบบเพียร์สัน (บุญชม ศรีสะอาด. 2532)

$$r = \frac{N \sum xy - (\sum x)(\sum y)}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}}$$

เมื่อ	r	แทน	ค่าอำนาจจำแนกรายข้อ
	N	แทน	จำนวนผู้เข้าสอบ
	X	แทน	คะแนนของข้อสอบข้อที่หาอำนาจจำแนก
	Y	แทน	คะแนนรวมของข้อสอบทุกข้อ

หมายเหตุ นำค่า r ที่คำนวณได้ไปทดสอบนัยสำคัญทางสถิติโดยเปิดตาราง ค่าวิกฤติของการแจกแจงสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (พวงรัตน์ ทวีรัตน์. 2531 : 304)

3. ค่าความยากและค่าอำนาจจำแนกรายข้อของแบบทดสอบเลือกตอบโดยใช้โปรแกรมTestqual
4. สถิติวิเคราะห์

4.1 หาค่าความเที่ยงตรงเชิงโครงสร้างแบบหลายลักษณะ-หลายวิธี โดยแยกเป็นความเที่ยงตรงเชิงเหมือน(Convergent Validity) และความเที่ยงตรงเชิงจำแนก (Discrimination Validity) (Allen and Yen. 1979 : 109-110 ; citing Cambell and Fiske. 1959.)

4.2 หาค่าความเที่ยงตรงเชิงโครงสร้าง ด้วยสูตรของ ไฮเซน และบอร์นสเตดท์ (Heise and Bohrnstedt, 1970 : 116)

$$\rho_{\tau_k s_k} = \frac{\sum_{i=1}^n s_i f_{ik}}{\sqrt{\sum_{i=1}^n \sum_{j=1}^n \text{Cov}(x_i, x_j)}}$$

เมื่อ	$\rho_{\tau_k s_k}$	แทน	ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบ
	s_i	แทน	ส่วนเบี่ยงเบนมาตรฐานของข้อคำถามข้อที่ i
	f_{ik}	แทน	น้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k
	$\text{Cov}(x_i, x_j)$	แทน	ค่าความแปรปรวนร่วมระหว่างคำถามข้อที่ i กับคำถามข้อที่ j

4.3 หาค่าความเชื่อมั่น ด้วยสูตร Ω_w ของอัลเลน (Allen, 1974)

$$\Omega_w = \frac{\sum_i \frac{\lambda_i^2}{1 - \lambda_i^2}}{1 + \sum_i \frac{\lambda_i^2}{1 - \lambda_i^2}}$$

เมื่อ	Ω_w	แทน	สัมประสิทธิ์ความเชื่อมั่นโอเมก้า แบบกำหนดน้ำหนักไม่เท่ากัน
	λ_i	แทน	น้ำหนักองค์ประกอบ(Factor Loading) ของคำถามข้อที่ i

4.4 หาค่าความเชื่อมั่นของแบบทดสอบ ด้วยสูตรของ Liou (บุญเชิด ภิญญอนันตพงษ์. 2537: 56 ; อ้างอิงจาก Liou, 1989)

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

เมื่อ	r_{L2}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ
	s_g^2	แทน	ความแปรปรวนของคะแนนในแต่ละตอน
	s_x^2	แทน	ความแปรปรวนของคะแนนสอบทั้งฉบับ
	λ_g^2	แทน	ผลรวมของความแปรปรวนร่วมระหว่างส่วนย่อยของแบบทดสอบ

บทที่ 4

ผลการวิเคราะห์ข้อมูล

สัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล

เพื่อให้เข้าใจตรงกันในการแปลความหมายผลการวิเคราะห์ข้อมูล ผู้วิจัยได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูล ดังต่อไปนี้

N	แทน	จำนวนนักเรียนที่เป็น กลุ่มตัวอย่าง
n	แทน	จำนวนข้อคำถามของแบบทดสอบ
k	แทน	คะแนนเต็มของแบบทดสอบ
$\bar{X}$	แทน	ค่าเฉลี่ย
S	แทน	ส่วนเบี่ยงเบนมาตรฐาน
p	แทน	ค่าความยากของข้อสอบ
r	แทน	ค่าอำนาจจำแนกของข้อสอบ
CV	แทน	สัมประสิทธิ์การกระจาย
OEQ	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด
OEQ -1	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิดด้านความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์
OEQ -2	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิดด้านการแก้โจทย์ปัญหา
OEQ -3	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิดด้านการให้เหตุผล
MCQ	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบ
MCQ -1	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบด้านความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์
MCQ -2	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบด้านการแก้โจทย์ปัญหา
MCQ -3	แทน	แบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบด้านการให้เหตุผล
MTMM	แทน	ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบที่หาด้วยวิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี(Multitrait-Multimethod)
$\rho_{T_{S_k}}$	แทน	ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบที่คำนวณด้วยสูตรของไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt.1970)
Ω_w	แทน	สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบที่คำนวณจากสูตรของอัลเลน (Allen.1974)
r_{L2}	แทน	สัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบที่คำนวณจากสูตรของเลียว (Liou . 1989)
$Cov(x_i, x_j)$	แทน	ความแปรปรวนร่วมระหว่างข้อคำถามระหว่างคำถามข้อที่ i กับคำถามข้อที่ j
S_i	แทน	ส่วนเบี่ยงเบนมาตรฐานของคำถามข้อที่ i
f_{ik}	แทน	น้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k

$S_i f_{ik}$	แทน	ผลคูณระหว่างส่วนเบี่ยงเบนมาตรฐานของคำถามข้อที่ i กับน้ำหนักองค์ประกอบของคำถามข้อที่ i จากองค์ประกอบที่ k
λ_i	แทน	น้ำหนักองค์ประกอบของคำถามข้อที่ i
λ_g^2	แทน	ผลรวมของความแปรปรวนร่วมระหว่างคะแนนในแต่ละด้านของแบบทดสอบ
S_g^2	แทน	ความแปรปรวนของคะแนนในแต่ละตอน
S_x^2	แทน	ความแปรปรวนของคะแนนสอบทั้งฉบับ

การนำเสนอผลการวิเคราะห์ข้อมูล

การศึกษาค้นคว้าครั้งนี้ได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับชั้น ดังต่อไปนี้

1. การประเมินคุณภาพของแบบทดสอบขั้นต้น ได้แก่การหาความเที่ยงตรงตามเนื้อหา

เกณฑ์การให้คะแนน

2. การหาค่าความยากและค่าอำนาจจำแนก
3. ค่าสถิติพื้นฐานของแบบทดสอบ
4. การแสดงหลักฐานค่าความเที่ยงตรงเชิงโครงสร้างแบบหลายลักษณะ-หลายวิธี
5. การแสดงหลักฐานค่าความเที่ยงตรงเชิงโครงสร้าง ที่คำนวณด้วยสูตรของไฮเซนและบอร์นสเตดท์

(Heise and Bohrnstedt.1970)

6. ค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณด้วยสูตรของ อัลเลน (Allen.1974)
7. ค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณด้วยสูตรของ เลี้ยว (Liou .1989)

ผลการวิเคราะห์ข้อมูล

1. การประเมินคุณภาพของแบบทดสอบขั้นต้น ได้แก่การหาค่าความเที่ยงตรงตามเนื้อหา

เกณฑ์การให้คะแนน

- 1.1 รายละเอียดของแบบทดสอบวิชาคณิตศาสตร์แบบปลายเปิด

แบบทดสอบปลายเปิดวิชาคณิตศาสตร์ เป็นแบบทดสอบที่ให้ผู้ตอบอธิบายหรือแสดงเหตุผลในการตอบ หรือแสดงวิธีหาคำตอบ โดยการเขียนตอบในกระดาษ ซึ่งวัดด้านต่าง ๆ 3 ด้าน ดังนี้

- | | |
|--|-------------|
| 1.1.1 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ | จำนวน 3 ข้อ |
| 1.1.2 การแก้ปัญหา | จำนวน 4 ข้อ |
| 1.1.3 การให้เหตุผล | จำนวน 3 ข้อ |

แบบทดสอบวิชาคณิตศาสตร์แบบเลือกตอบ เป็นการนำเอาข้อคำถามของแบบทดสอบปลายเปิดมาสร้างเป็นแบบทดสอบแบบเลือกตอบชนิด 5 ตัวเลือก วัดด้านต่าง ๆ 3 ด้าน ดังนี้

- | | |
|--|--------------|
| 1.1.4 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ | จำนวน 10 ข้อ |
| 1.1.5 การแก้ปัญหา | จำนวน 10 ข้อ |
| 1.1.6 การให้เหตุผล | จำนวน 10 ข้อ |

1.2 ความเที่ยงตรงเชิงเนื้อหา

การสร้างแบบทดสอบปลายเปิด เมื่อเขียนคำถามตามพฤติกรรมที่คาดหวัง จำนวน 17 ข้อแล้ว นำแบบทดสอบที่สร้างขึ้นไปทดสอบครั้งที่ 1 กับกลุ่มตัวอย่าง จำนวน 30 คน เพื่อตรวจสอบความชัดเจนของ ข้อคำถาม สร้างเกณฑ์การให้คะแนน แก่ไข ปรับปรุงแล้วนำแบบทดสอบที่ได้ไปให้ผู้เชี่ยวชาญจำนวน 5 ท่าน พิจารณาความเที่ยงตรงตามเนื้อหา และเกณฑ์การให้คะแนน แก่ไข ปรับปรุงตามข้อมูลที่ได้แล้วจัดทำเป็น แบบทดสอบนำไปสอบครั้งที่ 2 กับกลุ่มตัวอย่างจำนวน 50 คน เพื่อหาค่าความยากและค่าอำนาจจำแนก แล้ว คัดเลือกข้อที่มีค่าความยากตั้งแต่ .20 ถึง .80 และค่าอำนาจจำแนกที่มีนัยสำคัญทางสถิติ ไว้จำนวน 10 ข้อ

การสร้างแบบทดสอบเลือกตอบ นำข้อคำถามของแบบทดสอบปลายเปิดที่คัดเลือกไว้จำนวน 10 ข้อ มาสร้างเป็นแบบทดสอบแบบเลือกตอบ ชนิด 5 ตัวเลือก จำนวน 43 ข้อ นำไปให้ผู้เชี่ยวชาญชุดเดิม ตรวจสอบความเที่ยงตรงตามเนื้อหา แก่ไข ปรับปรุงจัดทำเป็นแบบทดสอบจำนวน 43 ข้อ แล้วนำไปสอบกับ กลุ่มตัวอย่างกลุ่มเดียวกันกับที่สอบแบบทดสอบปลายเปิดครั้งที่ 2 จำนวน 50 คน แล้วคัดเลือกข้อที่มีค่า ความยากในช่วง .20 ถึง .80 และค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป จำนวน 30 ข้อ

2. การหาค่าความยากและค่าอำนาจจำแนก

2.1 ค่าสถิติพื้นฐาน

ผู้วิจัยได้นำแบบทดสอบ ที่ผ่านการคัดเลือกคุณภาพด้านความเที่ยงตรงตามเนื้อหาทั้ง 2 ฉบับ ไปสอบกับกลุ่มตัวอย่างจำนวน 50 คน แล้วนำคะแนนที่ได้มาหาค่าสถิติพื้นฐาน ได้แก่ ค่าเฉลี่ยของแบบ ทดสอบ และส่วนเบี่ยงเบนมาตรฐาน ได้ผลการวิเคราะห์ดังตาราง 5

ตาราง 5 ค่าสถิติพื้นฐานของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด และฉบับเลือกตอบ(N = 50)

แบบทดสอบ	N	n	k	$\bar{X}$	S
OEQ	50	17	68	34.60	9.51
MCQ	50	43	43	21.20	7.41

ผลการวิเคราะห์ตามตาราง 5 ปรากฏว่า แบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด มีคะแนน เฉลี่ยเท่ากับ 34.60 ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 9.51 และแบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบ มี คะแนนเฉลี่ยเท่ากับ 21.20 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 7.41

2.2 การวิเคราะห์ข้อสอบรายข้อ

2.2.1 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด

ผู้วิจัยนำคะแนนที่ได้จากการสอบครั้งที่ 2 ของแบบทดสอบวิชาคณิตศาสตร์ ฉบับ ปลายเปิด จำนวน 17 ข้อ ผู้เข้าสอบจำนวน 50 คนมาวิเคราะห์หาค่าความยากและค่าอำนาจจำแนก ได้ผล การวิเคราะห์ดังตาราง 6

ตาราง 6 ค่าความยาก(p) และค่าอำนาจจำแนก (r) ของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด

ข้อที่	ค่าความยาก (p)	ค่าอำนาจจำแนก(r)	หมายเหตุ
1	.56	.34*	เลือกไว้
2	.45	.26	คัดออก
3	.14	.54*	คัดออก
4	.46	.32*	เลือกไว้
5	.47	.36*	เลือกไว้
6	.54	.34*	เลือกไว้
7	.53	.31*	เลือกไว้
8	.72	.52*	เลือกไว้
9	.56	.49*	เลือกไว้
10	.62	.59*	เลือกไว้
11	.58	.45*	เลือกไว้
12	.58	.46*	เลือกไว้
13	.59	.51*	เลือกไว้
14	.50	.55*	เลือกไว้
15	.58	.53*	เลือกไว้
16	.44	.64*	เลือกไว้
17	.54	.45*	เลือกไว้

*มีนัยสำคัญทางสถิติที่ระดับ .05

ผลการวิเคราะห์ตามตาราง 6 ปรากฏว่า แบบทดสอบปลายเปิดวิชาคณิตศาสตร์ จำนวน 17 ข้อ มีค่าความยากตั้งแต่ .14 ถึง .74 และมีค่าอำนาจจำแนกตั้งแต่ .23 ถึง .64 มีข้อคำถามที่ไม่มีนัยสำคัญทางสถิติ 1 ข้อ คือ ข้อ 2 ผู้วิจัยจึงตัดข้อที่มีค่าความยากน้อยกว่า .20 และข้อที่ค่าอำนาจจำแนกไม่มีนัยสำคัญทางสถิติที่ระดับ .05 ออก หลังจากตัดข้อคำถามดังกล่าวออกแล้ว แบบทดสอบยังมีจำนวนข้อมาก ดังนั้น ผู้วิจัยจึงคัดเลือกข้อที่มีค่าความยากตั้งแต่ .46 ถึง .62 และค่าอำนาจจำแนกตั้งแต่ .31 ถึง .59 มาใช้ในการทดลองสอบครั้งที่ 3 จำนวน 10 ข้อ แบ่งเป็น 3 ด้าน ดังนี้

ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ จำนวน 3 ข้อ

การแก้ปัญหา จำนวน 4 ข้อ

การให้เหตุผล จำนวน 3 ข้อ

2.2.2 ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบ ผู้วิจัยนำคะแนนที่ได้จากการสอบแบบทดสอบวิชาคณิตศาสตร์ ฉบับเลือกตอบ จำนวน 43 ข้อ ผู้เข้าสอบจำนวน 50 คน มาวิเคราะห์หาค่าความยากและค่าอำนาจจำแนก โดยใช้โปรแกรม Testqual ได้ผลการวิเคราะห์ดังตาราง 7

ตาราง 7 ค่าความยาก(p) และค่าอำนาจจำแนก(r) ของแบบทดสอบวิชาคณิตศาสตร์ ฉบับเลือกตอบ

ข้อที่	ค่าความยาก(p)	ค่าอำนาจจำแนก(r)	หมายเหตุ
1	.26	.20	เลือกไว้
2	.88	.16	คัดออก
3	.46	.20	เลือกไว้
4	.40	.24	เลือกไว้
5	.72	.16	คัดออก
6	.58	.12	คัดออก
7	.86	.28	เลือกไว้
8	.38	.36	เลือกไว้
9	.72	.40	เลือกไว้
10	.60	.24	เลือกไว้
11	.48	.24	เลือกไว้
12	.38	.04	คัดออก
13	.56	.00	คัดออก
14	.30	.12	คัดออก
15	.60	.32	เลือกไว้
16	.68	.24	เลือกไว้
17	.58	.28	เลือกไว้
18	.36	.40	เลือกไว้
19	.32	.24	เลือกไว้
20	.24	.16	คัดออก
21	.40	.24	เลือกไว้
22	.52	.24	เลือกไว้
23	.32	.24	เลือกไว้
24	.46	.20	เลือกไว้
25	.70	.51	เลือกไว้
26	.38	.20	เลือกไว้
27	.18	.28	คัดออก
28	.04	.00	คัดออก
29	.54	.28	เลือกไว้
30	.56	.48	เลือกไว้
31	.70	.59	เลือกไว้
32	.14	.20	คัดออก
33	.26	.12	คัดออก
34	.38	.36	เลือกไว้
35	.58	.51	เลือกไว้

ตาราง 7 (ต่อ)

ข้อที่	ค่าความยาก(p)	ค่าอำนาจจำแนก(r)	หมายเหตุ
36	.62	.59	เลือกไว้
37	.68	.48	เลือกไว้
38	.64	.63	เลือกไว้
39	.60	.55	เลือกไว้
40	.62	.44	เลือกไว้
41	.66	.51	เลือกไว้
42	.46	.28	เลือกไว้
43	.42	.28	เลือกไว้

ผลการวิเคราะห์ตามตาราง 7 ปรากฏว่า แบบทดสอบวิชาคณิตศาสตร์ ฉบับเลือกตอบ มีค่าความยากตั้งแต่ .04 ถึง .88 ค่าอำนาจจำแนกตั้งแต่ .00 ถึง .63 ผู้วิจัยจึงตัดข้อที่มีค่าความยากน้อยกว่า .20 และอำนาจจำแนกน้อยกว่า .20 ออก แล้วคัดเลือกข้อที่มีค่าความยากตั้งแต่ .26 ถึง .72 และค่าอำนาจจำแนกตั้งแต่ .20 ถึง .63 มาใช้ในการทดลองสอบครั้งที่ 2 จำนวน 30 ข้อ แบ่งเป็น 3 ด้าน ดังนี้

ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ จำนวน 10 ข้อ

การแก้ปัญหา จำนวน 10 ข้อ

การให้เหตุผล จำนวน 10 ข้อ

3. ค่าสถิติพื้นฐานของแบบทดสอบ

ผู้วิจัยนำคะแนนที่ได้จากการสอบกลุ่มตัวอย่างจำนวน 300 คน ของแบบทดสอบทั้งสองฉบับ มาวิเคราะห์หาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสัมประสิทธิ์การกระจาย โดยแยกเป็นรายด้านและรวมทั้งฉบับ ได้ผลการวิเคราะห์ดังตาราง 8

ตาราง 8 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และสัมประสิทธิ์การกระจาย ของคะแนนแต่ละด้าน และคะแนนรวมของแบบทดสอบแต่ละฉบับ(N = 300)

แบบทดสอบ	n	k	$\bar{X}$	S	CV
OEQ -1	3	12	4.6167	2.4130	0.5227
OEQ -2	4	16	9.0933	2.8540	0.3139
OEQ -3	3	12	7.1100	2.5297	0.3558
OEQ	10	40	20.8200	5.6433	0.2711
MCQ -1	10	10	5.6667	1.7104	0.3018
MCQ -2	10	10	7.1600	2.1564	0.3012
MCQ -3	10	10	6.2633	1.8526	0.2958
MCQ	30	30	19.0900	4.2250	0.2213

ผลการวิเคราะห์ตามตาราง 8 ปรากฏว่า ค่าเฉลี่ยของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด มีค่าเท่ากับ 20.8200 โดยด้านการแก้ปัญหาที่มีค่าเฉลี่ยสูงสุด รองลงมาคือด้านการให้เหตุผล และด้านความคิดรวบยอดทางคณิตศาสตร์ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 5.6433 และค่าสัมประสิทธิ์การกระจายมีค่าเท่ากับ 0.2711 โดยที่ด้านความคิดรวบยอดทางคณิตศาสตร์ มีการกระจายมากที่สุด ส่วนแบบทดสอบวิชาคณิตศาสตร์ ฉบับเลือกตอบมีค่าเฉลี่ยเท่ากับ 19.0900 โดยด้านการแก้ปัญหาที่มีค่าเฉลี่ยสูงสุด รองลงมาคือด้านการให้เหตุผล และด้านความคิดรวบยอดทางคณิตศาสตร์ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 4.2255 และค่าสัมประสิทธิ์การกระจาย มีค่าเท่ากับ 0.2213 โดยที่ด้านความคิดรวบยอดทางคณิตศาสตร์ มีการกระจายมากที่สุด

4. การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างแบบหลายลักษณะ-หลายวิธี (Multitrait-Multimethod)

ผู้วิจัยนำคะแนนที่ได้จากการสอบครั้งที่ 3 ของแบบทดสอบวิชาคณิตศาสตร์ฉบับปลายเปิด และจากการสอบครั้งที่ 2 ของแบบทดสอบวิชาคณิตศาสตร์ฉบับเลือกตอบ มาหาค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน เพื่อนำค่าที่ได้ไปหาค่าความเที่ยงตรงเชิงโครงสร้างแบบหลายลักษณะ-หลายวิธี (Multitrait-Multimethod) โดยแยกเป็นความเที่ยงตรงเชิงเหมือน (Convergent Validity) และความเที่ยงตรงเชิงจำแนก (Discriminant Validity) ได้ผลการวิเคราะห์ดังตาราง 9

ตาราง 9 ค่าความเที่ยงตรงเชิงโครงสร้างแบบหลายลักษณะ-หลายวิธีของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์

	OEQ -1	OEQ -2	OEQ -3	MCQ -1	MCQ -2	MCQ -3
OEQ -1	.7316					
OEQ -2	(.3062**)	.7032				
OEQ -3	(.2398**)	(.3048**)	.7037			
MCQ -1	.4959**	.3272**	.2488**	.7228		
MCQ -2	.2832**	.5361**	.4753**	(.2884**)	.6728	
MCQ -3	.1508**	.3056**	.4862**	(.1863**)	(.4150**)	.7184

**มีนัยสำคัญทางสถิติที่ระดับ .01

ตัวเลขในแนวทแยงมุมหลักเป็นค่าความเชื่อมั่น

ตัวเลขที่ขีดเส้นใต้ เป็นความเที่ยงตรงเชิงเหมือน

ตัวเลขที่อยู่ในวงเล็บเป็นความเที่ยงตรงเชิงจำแนก ที่วัดลักษณะต่างกัน ด้วยแบบทดสอบชนิดเดียวกัน

ตัวเลขที่เหลือเป็นความเที่ยงตรงเชิงจำแนก ที่วัดลักษณะต่างกัน ด้วยแบบทดสอบต่างชนิดกัน

ผลการวิเคราะห์ตามตาราง 9 ปรากฏว่า ค่าความเที่ยงตรงเชิงโครงสร้างแบบหลายลักษณะ-หลายวิธี (Multitrait-Multimethod) ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ มีค่าสัมประสิทธิ์ความเชื่อมั่น ซึ่งเป็นการวัดในลักษณะเดียวกันด้วยแบบทดสอบชนิดเดียวกัน จะมีค่าสหสัมพันธ์กันสูงสุด คือ มีค่าตั้งแต่ .6728 ถึง .7316 สัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงเหมือน (Convergent Validity) ซึ่งเป็นการวัดในลักษณะเดียวกันด้วยแบบทดสอบต่างชนิดกัน มีค่าตั้งแต่ .4862 ถึง .5361 (ตัวเลขที่ขีดเส้นใต้) ซึ่งมีค่าสหสัมพันธ์กันต่ำกว่าค่าสัมประสิทธิ์ความเชื่อมั่น โดยที่แบบทดสอบวิชาคณิตศาสตร์ด้านการแก้ปัญหา มีค่า

สัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงเหมือนสูงสุด และแบบทดสอบวิชาคณิตศาสตร์ด้านการให้เหตุผล มีค่าสัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงเหมือนต่ำสุด ส่วนค่าสัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรงเชิงจำแนก (Discriminant Validity) ที่เป็นการวัดลักษณะต่างกันด้วยแบบทดสอบชนิดเดียวกัน มีค่าตั้งแต่ .1863 ถึง .4150 และ ที่เป็นการวัดลักษณะต่างกันด้วยแบบทดสอบต่างชนิดกันมีค่าตั้งแต่ .1508 ถึง .4753

5. การแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้าง ที่คำนวณด้วยสูตรของไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt.1970.)

ผู้วิจัยนำคะแนนที่ได้จากการสอบแบบทดสอบปลายเปิด ครั้งที่ 3 จำนวน 10 ข้อ ผู้เข้าสอบจำนวน 300 คน มาหาค่าความแปรปรวนร่วมระหว่างข้อคำถาม ส่วนเบี่ยงเบนมาตรฐานรายข้อ และหาหน้าหนักองค์ประกอบของข้อคำถามแต่ละข้อจากองค์ประกอบที่ k โดยใช้โปรแกรม SPSS FOR WINDOWS แล้วนำค่าดังกล่าวที่ได้ไปแทนในสูตร $\rho_{T_k S_k}$ ของไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt.1970.) เพื่อหาค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ได้ผลการวิเคราะห์ดังตาราง 10 , 11 และ 12

ตาราง 10 ค่าความแปรปรวนร่วมระหว่างข้อคำถาม ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์
จำนวน 10 ข้อ

ข้อ	1	2	3	4	5	6	7	8	9	10	$\Sigma Cov(x_i, x_j)$
1	1.3144										1.3144
2	.2767	1.6620									2.1387
3	.1176	.0858	1.9296								2.1330
4	.1598	-.0484	-.1272	1.6352							1.9706
5	-.0522	.0665	-.1086	.0592	1.7882						2.0747
6	.3866	.2269	.3440	.2018	.1931	1.6675					3.0199
7	.3146	.2076	.2051	.0742	-.0205	.3473	1.6591				2.8284
8	.1125	-.0915	.1758	.1157	.1540	.3502	.1460	2.0506			3.1963
9	.0764	.0798	.0761	.1588	-.0745	.1470	.1021	.1611	1.0687		1.9445
10	.2507	.2499	.2374	.1979	.3643	.4805	.2376	.3221	.2879	2.4162	5.0445
$\Sigma \Sigma Cov(x_i, x_j)$											25.6650

ผลการวิเคราะห์ตามตาราง 10 เป็นการนำเสนอค่าความแปรปรวนร่วมระหว่างข้อคำถาม จำนวน 10 ข้อ ได้ผลรวมของความแปรปรวนร่วมทั้งหมด เท่ากับ 25.6650 นั่นคือ $\Sigma \Sigma Cov(x_i, x_j) = 25.6650$ แล้วนำค่า $\Sigma \Sigma Cov(x_i, x_j)$ ที่ได้ไปแทนในสูตร $\rho_{T_k S_k}$ ของไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt.1970.) เพื่อหาค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ในตาราง 12

ตาราง 11 ส่วนเบี่ยงเบนมาตรฐานรายข้อ(S_i) น้ำหนักองค์ประกอบของคำถามแต่ละข้อในองค์ประกอบที่ k (f_{ik}) และผลคูณระหว่างส่วนเบี่ยงเบนมาตรฐานรายข้อ กับน้ำหนักองค์ประกอบของคำถามแต่ละข้อในองค์ประกอบที่ k ($S_i f_{ik}$)

ข้อ	S_i	f_{ik}	$S_i f_{ik}$
1	1.1465	0.6179	0.7084
2	1.3646	0.5291	0.7220
3	1.3891	0.4772	0.6629
4	1.2787	0.0588	0.0752
5	1.3372	0.1738	0.2324
6	1.2913	0.5335	0.6889
7	1.2881	0.5869	0.7559
8	1.4320	0.0650	0.0931
9	1.0338	0.1702	0.1760
10	1.5544	0.2839	.04413
$\Sigma S_i f_{ik}$			4.5560

ผลการวิเคราะห์ตามตาราง 11 ปรากฏว่า ผลรวมของผลคูณระหว่างส่วนเบี่ยงเบนมาตรฐานรายข้อ กับน้ำหนักองค์ประกอบของคำถามแต่ละข้อ ในองค์ประกอบที่ k มีค่าเท่ากับ 4.5560 นั่นคือ $\Sigma S_i f_{ik} = 4.5560$ แล้วนำค่า $\Sigma S_i f_{ik}$ ที่ได้ไปแทนในสูตร $\rho_{T_k S_k}$ ของไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt. 1970.) เพื่อหาค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ในตาราง 12

ตาราง 12 ค่าความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร $\rho_{T_k S_k}$ ของ ไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt. 1970.)

แบบทดสอบ	$\Sigma S_i f_{ik}$	$\Sigma \Sigma \text{Cov}(x_i, x_j)$	$\rho_{T_k S_k}$
OEQ	4.5560	25.6550	0.8993

ผลการวิเคราะห์ตามตาราง 12 ปรากฏว่า เมื่อนำค่า $\Sigma S_i f_{ik}$ และ $\Sigma \Sigma \text{Cov}(x_i, x_j)$ ที่ได้ไปแทนในสูตร $\rho_{T_k S_k}$ ของไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt. 1970.) เพื่อหาค่าความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ หลังจากคำนวณแล้วได้ค่า $\rho_{T_k S_k} = 0.8993$ ซึ่งสอดคล้องกับสมมุติฐานข้อ 2

6. การหาค่าความเชื่อมั่นของแบบทดสอบ ที่คำนวณด้วยสูตรของอัลเลน (Allen.1974 : Ω_w)

ผู้วิจัยนำคะแนนที่ได้จากการสอบแบบทดสอบปลายเปิด ครั้งที่ 3 มาวิเคราะห์หาค่าน้ำหนัก

องค์ประกอบ (Factor Loading : λ_i) โดยใช้โปรแกรม SPSS FOR WINDOWS เพื่อนำค่าที่ได้ไปหาค่ากำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ (λ_i^2) ผลต่างระหว่าง 1 กับกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ ($1-\lambda_i^2$) และสัดส่วนของ กำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อต่อผลต่างระหว่าง 1 กับกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ ($\lambda_i^2 / 1-\lambda_i^2$) แล้วนำค่าดังกล่าวที่ได้ มาแทนในสูตร Ω_w ของอัลเลน (Allen.1974) เพื่อหาค่าความเชื่อมั่นของแบบทดสอบของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ได้ผลการวิเคราะห์ดังตาราง 13 และ 14

ตาราง 13 กำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ (λ_i^2) ผลต่างระหว่าง 1 กับกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ ($1-\lambda_i^2$) และสัดส่วนของกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อต่อผลต่างระหว่าง 1 กับกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อ ($\lambda_i^2 / 1-\lambda_i^2$)

ข้อ	λ_i	λ_i^2	$1-\lambda_i^2$	$\lambda_i^2 / 1-\lambda_i^2$
1	0.6179	0.3818	0.6182	0.6175
2	0.5291	0.2799	0.7201	0.3887
3	0.4772	0.2278	0.7721	0.2949
4	0.6267	0.3927	0.6073	0.6466
5	0.8607	0.7408	0.2592	2.8583
6	0.5335	0.2846	0.7154	2.3978
7	0.5869	0.3444	0.6556	0.5254
8	0.4924	0.2426	0.7576	0.3200
9	0.6140	0.3770	0.7230	0.6052
10	0.5242	0.2747	0.7253	0.3788
$\Sigma \lambda_i^2 / 1-\lambda_i^2$				7.0333

ผลการวิเคราะห์ตามตาราง 13 ปรากฏว่าผลรวมของสัดส่วนของกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อต่อผลต่างระหว่าง 1 กับกำลังสองของน้ำหนักองค์ประกอบของข้อคำถามในแต่ละข้อมีค่าเท่ากับ 7.0333 นั่นคือ $\Sigma \lambda_i^2 / 1-\lambda_i^2 = 7.0333$ แล้วนำค่า $\lambda_i^2 / 1-\lambda_i^2$ ไปแทนในสูตร Ω_w ของอัลเลน (Allen.1974) เพื่อหาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ในตาราง 14

ตาราง 14 ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของอัลเลน
(Allen.1974. : Ω_w)

แบบทดสอบ	$\Sigma \lambda_i^2 / 1 - \lambda_i^2$	$1 + \Sigma \lambda_i^2 / 1 - \lambda_i^2$	Ω_w
OEQ	7.0333	8.0333	0.8755

ผลการวิเคราะห์ตามตาราง 14 ปรากฏว่าเมื่อนำค่า $\Sigma \lambda_i^2 / 1 - \lambda_i^2$ และ $1 + \Sigma \lambda_i^2 / 1 - \lambda_i^2$ ที่ได้ไปแทนในสูตร Ω_w ของอัลเลน (Allen.1974) เพื่อหาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ หลังจากคำนวณแล้ว ได้ค่า $\Omega_w = 0.8755$ ซึ่งสอดคล้องกับสมมุติฐานข้อ 3

7. การหาค่าความเชื่อมั่นของ แบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของเลียว (Liou.1989 : r_{L2})

ผู้วิจัยนำคะแนนที่ได้จากการสอบแบบทดสอบปลายเปิดครั้งที่ 3 มาหาค่าความแปรปรวนของคะแนนสอบทั้งฉบับ (S_g^2) ความแปรปรวนของคะแนนในแต่ละด้าน (S_g^2) และความแปรปรวนร่วมระหว่างคะแนนในแต่ละด้านของแบบทดสอบ (λ_g^2) แล้วนำค่าดังกล่าวที่ได้ไปแทนในสูตร r_{L2} ของเลียว (Liou.1989) เพื่อ หาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ได้ผลการวิเคราะห์ดังตาราง 15 และ 16

ตาราง 15 ค่าความแปรปรวนร่วมระหว่างคะแนนในแต่ละด้าน (λ_g) ผลรวมของความแปรปรวนร่วมระหว่างคะแนนในแต่ละด้าน (λ_g^2) และความแปรปรวนของคะแนนในแต่ละด้าน (S_g^2)

แบบทดสอบ	λ_g	λ_g^2	S_g^2
OEQ -1	0.2932	0.0860	5.8225
OEQ -2	0.3263	0.1065	8.1451
OEQ -3	0.2739	0.0750	2.5297
Σ		0.2675	16.4973

ผลการวิเคราะห์ตามตาราง 15 ปรากฏว่า ผลรวมของความแปรปรวนร่วมระหว่างคะแนนในแต่ละด้าน มีค่าเท่ากับ 0.2675 นั่นคือ $\lambda_g^2 = 0.2675$ และผลรวมของความแปรปรวนของคะแนนในแต่ละด้าน มีค่าเท่ากับ 16.4973 นั่นคือ $S_g^2 = 16.4973$ แล้วนำค่าดังกล่าวที่ได้ ไปแทนในสูตร r_{L2} ของเลียว (Liou.1989) เพื่อ หาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ในตาราง 16

ตาราง 16 ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของเลียว
(Liou.1989 : r_{L2})

แบบทดสอบ	$\sum \lambda_g^2$	$\sum s_g^2$	S^2	r_{L2}
OEQ	0.2675	16.4973	31.8471	0.6580

ผลการวิเคราะห์ตามตาราง 16 ปรากฏว่า เมื่อนำค่า $\sum \lambda_g^2$, $\sum s_g^2$, S^2 ที่ได้มาแทนค่าในสูตร r_{L2} ของเลียว(Liou.1989) เพื่อหาความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ หลังจากคำนวณแล้วได้ค่า $r_{L2} = 0.6580$ ซึ่งสอดคล้องกับสมมุติฐานข้อ 4

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ความมุ่งหมายของการศึกษาค้นคว้า

1. เพื่อแสดงหลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ สองวิธี ดังนี้
 - 1.1 วิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี(Multitrait-Multimethod)
 - 1.2 วิธีหาความสัมพันธ์ระหว่างคะแนนจริงกับคะแนนองค์ประกอบ ที่คำนวณด้วยสูตร P_{TS_K} ของไอเซน และบอร์นสเตดท์ (Heise and Bohmstedt)
2. เพื่อหาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ สองวิธี ดังนี้
 - 2.1 วิธีวิเคราะห์น้ำหนักองค์ประกอบ ที่คำนวณด้วยสูตร Ω_w ของอัลเลน (Allen.1974)
 - 2.2 วิธีวิเคราะห์แบบคะแนนจริงสัมพันธ์ ที่คำนวณด้วยสูตร r_{L2} ของเลียว (Liou.1989)

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 ของโรงเรียนในสังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพมหานคร ที่เป็นโรงเรียนสหศึกษา จำนวน 11 โรงเรียน 159 ห้องเรียน จำนวน 6,987 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 ของโรงเรียนในสังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพมหานคร ที่เป็นโรงเรียนสหศึกษา จำนวน 300 คน

เครื่องมือที่ใช้ในการศึกษาค้นคว้า

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ ประกอบด้วย

1. แบบทดสอบปลายเปิด จำนวน 10 ข้อ ข้อละ 4 คะแนน แบ่งเป็น 3 ด้าน คือ
 - 1.1 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ จำนวน 3 ข้อ
 - 1.2 การแก้ปัญหา จำนวน 4 ข้อ
 - 1.3 การให้เหตุผล จำนวน 3 ข้อ
2. แบบทดสอบเลือกตอบ จำนวน 30 ข้อ ข้อละ 1 คะแนน แบ่งเป็น 3 ด้าน คือ
 - 2.1 ความเข้าใจในความคิดรวบยอดทางคณิตศาสตร์ จำนวน 10 ข้อ
 - 2.2 การแก้ปัญหา จำนวน 10 ข้อ
 - 2.3 การให้เหตุผล จำนวน 10 ข้อ

วิธีดำเนินการเก็บรวบรวมข้อมูล

การดำเนินการเก็บรวบรวมข้อมูล ผู้วิจัยได้ดำเนินการตามลำดับขั้นดังต่อไปนี้

1. ทำหนังสือขอความร่วมมือ จากผู้อำนวยการโรงเรียนที่เป็นกลุ่มทดลองเครื่องมือและกลุ่มตัวอย่างเพื่อกำหนดวัน เวลา และสถานที่
2. ชี้แจงให้นักเรียนทราบถึงวัตถุประสงค์ของการสอบ และขอความร่วมมือในการสอบ เพื่อให้ได้ผลตามความเป็นจริง
3. นักเรียนทั้งกลุ่มทดลองเครื่องมือและกลุ่มตัวอย่างจะสอบด้วยข้อสอบทั้ง 2 ฉบับ โดยสอบด้วยแบบทดสอบปลายเปิดก่อน หลังจากนั้นประมาณ 2 สัปดาห์ จึงสอบด้วยแบบทดสอบเลือกตอบ
4. ตรวจให้คะแนนตามเกณฑ์ที่กำหนดไว้
5. นำคะแนนที่ได้มาวิเคราะห์หาค่าสถิติ

สรุปผลการวิเคราะห์ข้อมูล

1. ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่ใช้วิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) พบว่า สัมประสิทธิ์ค่าความเชื่อมั่น ซึ่งเป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จาก การวัดลักษณะเดียวกันด้วยแบบทดสอบชนิดเดียวกัน จะมีค่าสูงสุด คือ มีค่าตั้งแต่ .6728 ถึง .7316 ความเที่ยงตรงเชิงเหมือน ซึ่งเป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จาก การวัดลักษณะเดียวกันด้วยแบบทดสอบต่างชนิดกัน จะมีค่าต่ำกว่าสัมประสิทธิ์ความเชื่อมั่น คือมีค่าตั้งแต่ .4862 ถึง .5361 และความเที่ยงตรงเชิงจำแนก ที่เป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จาก การวัดลักษณะต่างกันด้วยแบบทดสอบชนิดเดียวกัน มีค่าตั้งแต่ .1863 ถึง .4150 และที่เป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จาก การวัดลักษณะต่างกันด้วยแบบทดสอบต่างชนิดเดียวกัน มีค่าตั้งแต่ .1508 ถึง .4753
2. ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร ρ_{T_k} ของไฮเซนและบรอนสเตดท์ (Heise and Bohrnstedt . 1970) ซึ่งเป็นการหาความสัมพันธ์ระหว่างคะแนนจริงของคุณลักษณะที่ได้ (T_k) กับคะแนนองค์ประกอบ (S_k) มีค่าเท่ากับ .8993
3. ความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร Ω_u ของอัลเลน (Allen.1974) ซึ่งเป็นการหาค่าความเชื่อมั่นที่ให้ความสำคัญกับค่าน้ำหนักองค์ประกอบ (Factor Loading) แบบกำหนดน้ำหนักไม่เท่ากัน มีค่าเท่ากับ .8755
4. ความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร r_{L2} ของเลียว (Liou.1989) ซึ่งเป็นการหาค่าความเชื่อมั่นแบบคะแนนจริงสัมพันธ์ (Congeneric) มีค่าเท่ากับ .6580

0.26 - 0.77

0.25 - 0.84

อภิปรายผล

จากการศึกษาครั้งนี้ มีจุดมุ่งหมายที่สำคัญคือ ต้องการศึกษาคูณภาพของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ โดยใช้การแสดงผลหลักฐานความเที่ยงตรงเชิงโครงสร้าง และค่าความเชื่อมั่นของแบบทดสอบ ซึ่งอภิปรายผลได้ดังนี้

1. แบบทดสอบปลายเปิดวิชาคณิตศาสตร์ มีค่าความยาก ตั้งแต่ .46 ถึง .62 ซึ่งค่าความยากของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์แต่ละข้อ มีค่าความยาก อยู่ในช่วง .20 ถึง .80 (ชวาล แพทย์กุล. 2516 : 274) และมีค่าอำนาจจำแนกตั้งแต่ .31 ถึง .59 เมื่อนำค่าอำนาจจำแนกที่ได้ไปทดสอบนัยสำคัญทางสถิติ โดยเปิดตารางค่าวิกฤติของการแจกแจงแบบเพียร์สัน (พวงรัตน์ ทวีรัตน์. 2531 : 304) ที่ $df = 50$, $\alpha = .05$ มีค่าเท่ากับ .2732 พบว่า ค่าอำนาจจำแนกทุกค่าที่คำนวณได้ มีค่าสูงกว่า .2732 แสดงว่าค่าอำนาจจำแนกทุกค่า ของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ มีนัยสำคัญทางสถิติ ยืนยันให้เห็นว่าแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ มีคุณภาพทั้งค่าความยาก และค่าอำนาจจำแนก ในระดับที่ใช้ประเมินผลความสามารถของผู้เข้าสอบได้ เช่นเดียวกับแบบทดสอบปรนัยแบบเลือกตอบ แต่ให้ความโดดเด่นที่สามารถวัดความคิดของผู้สอบได้ตรงกว่า เพราะไม่ต้องมีตัวเลือกให้ผู้เข้าสอบ ซึ่งเป็นเสมือนตัวแฉะความคิดของผู้เข้าสอบ ซึ่งสอดคล้องกับคำกล่าวของ กรอนลันด์ (Gronlund. 1965 : 108) ที่กล่าวว่า แบบทดสอบอัตนัยใช้ได้ดีกับการวัดความเข้าใจ ทักษะการคิด ความสามารถในการเขียน การแก้ปัญหา โดยเฉพาะการวัดความคิดริเริ่มสร้างสรรค์ และเหมาะสมอย่างยิ่งกับคำถามที่เน้นให้นักเรียนใช้ความสามารถในการหาคำตอบเอง

2. ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่ใช้วิธีหาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) พบว่า แบบทดสอบปลายเปิดวิชาคณิตศาสตร์ มีความเที่ยงตรงเชิงโครงสร้าง เป็นไปตามเกณฑ์ที่ แลมป์เบลและฟิสต์ (Allen and Yen.1979 : 109-110 ; citing Cambell and Fiske. 1959) กำหนดไว้คือ สัมประสิทธิ์ความเชื่อมั่น ซึ่งเป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดลักษณะเดียวกันด้วยแบบทดสอบชนิดเดียวกัน ($r(OEQ-1OEQ-1)$, $r(OEQ-2 OEQ-2)$, $r(OEQ-3OEQ-3)$, $r(MCQ-1 MCQ-1)$, $r(MCQ-2 MCQ-2)$, $r(MCQ-3 MCQ-3)$) จะมีค่าสูงสุด คือมีค่าตั้งแต่ .6728 ถึง .7316 ค่าความเที่ยงตรงเชิงเหมือน ซึ่งเป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดลักษณะเดียวกันด้วยแบบทดสอบต่างชนิดกัน ($r(OEQ-1 MCQ-1)$, $r(OEQ-2 MCQ-2)$, $r(OEQ-3 MCQ-3)$) จะมีค่าต่ำกว่าสัมประสิทธิ์ค่าความเชื่อมั่น คือมีค่าตั้งแต่ .4862 ถึง .5361 ทุกค่ามีนัยสำคัญทางสถิติที่ระดับ .01 และค่าความเที่ยงตรงเชิงจำแนก ซึ่งเป็นค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดลักษณะต่างกัน ด้วยแบบทดสอบชนิดเดียวกันหรือต่างชนิดกัน ($r(OEQ-1 OEQ-2)$, $r(OEQ-1 OEQ-3)$, $r(OEQ-2 OEQ-3)$, $r(MCQ-1 MCQ-2)$, $r(MCQ-1 MCQ-3)$, $r(MCQ-2 MCQ-3)$, $r(OEQ-1 MCQ-2)$, $r(OEQ-1 MCQ-3)$, $r(OEQ-2 MCQ-3)$, $r(OEQ-2 MCQ-1)$, $r(OEQ-3 MCQ-1)$, $r(OEQ-3 MCQ-2)$) จะมีค่าต่ำกว่าค่าความเที่ยงตรงเชิงเหมือน คือมีค่าตั้งแต่ .1508 ถึง .4753 และสัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดลักษณะต่างกัน ด้วยแบบทดสอบชนิดเดียวกัน ($r(OEQ-1 OEQ-2)$, $r(OEQ-1OEQ-3)$, $r(OEQ-2 OEQ-3)$, $r(MCQ-1 MCQ-2)$, $r(MCQ-1 MCQ-3)$, $r(MCQ-2 MCQ-3)$) ควรจะมีค่าสูงกว่า สัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดลักษณะต่างกันด้วยแบบทดสอบต่างชนิดกัน ($r(OEQ-1 MCQ-2)$, $r(OEQ-1 MCQ-3)$, $r(OEQ-2 MCQ-3)$, $r(OEQ-2 MCQ-1)$, $r(OEQ-3 MCQ-1)$, $r(OEQ-3 MCQ-2)$) ซึ่งผลจากการวิจัยครั้งนี้ มีบางค่าไม่เป็นไปตามเกณฑ์นี้ อย่างเช่น สัมประสิทธิ์สหสัมพันธ์ ที่ได้จากการวัดความคิดรวบยอดทางคณิตศาสตร์กับการให้เหตุผล เมื่อวัดด้วยแบบทดสอบฉบับเลือกตอบ ($r(MCQ-1 MCQ-3)$) ควรจะมีค่าสูงกว่า สัมประสิทธิ์สหสัมพันธ์ที่ได้จากการวัดความคิดรวบยอดทางคณิตศาสตร์ ที่สอบด้วยแบบทดสอบปลายเปิด กับ การแก้ปัญหาที่สอบด้วยแบบทดสอบเลือกตอบ ($r(OEQ-1 MCQ-2)$) สัมประสิทธิ์สหสัมพันธ์ ที่ได้จากการวัดการแก้ปัญหาที่สอบด้วยแบบทดสอบปลายเปิด กับ การ

ให้เหตุผลที่สอบด้วยแบบทดสอบเลือกตอบ ($r(\text{OEQ-2 MCQ-3})$) สัมประสิทธิ์สหสัมพันธ์ ที่ได้จากการวัดการแก้ปัญหาที่สอบด้วยแบบทดสอบปลายเปิด กับการวัดความคิดรวบยอดทางคณิตศาสตร์ที่สอบด้วยแบบทดสอบเลือกตอบ ($r(\text{OEQ-2 MCQ-1})$) สัมประสิทธิ์สหสัมพันธ์ ที่ได้จากการวัดการให้เหตุผลที่สอบด้วยแบบทดสอบปลายเปิด กับการวัดความคิดรวบยอดทางคณิตศาสตร์ ที่สอบด้วยแบบทดสอบเลือกตอบ ($r(\text{OEQ-3 MCQ-1})$) และสัมประสิทธิ์สหสัมพันธ์ ที่ได้จากการวัดการให้เหตุผลที่สอบด้วยแบบทดสอบปลายเปิด กับการแก้ปัญหาที่สอบด้วยแบบทดสอบเลือกตอบ ($r(\text{OEQ-3 MCQ-2})$) แต่ในที่นี้กลับมีค่าต่ำกว่า ทั้งนี้ อาจเนื่องมาจาก ผลการวิเคราะห์องค์ประกอบแบบทดสอบปลายเปิดจำนวน 10 ข้อ ปรากฏว่าได้ 4 องค์ประกอบแต่นิยามโครงสร้างการวัดไว้ 3 องค์ประกอบ และแบบทดสอบเลือกตอบ จำนวน 30 ข้อ ปรากฏว่าได้ 6 องค์ประกอบ แต่นิยามโครงสร้างการวัดไว้ 3 องค์ประกอบเช่นกัน อาจเป็นไปได้ว่าแต่ละด้านที่กำหนดไว้นั้นมีลักษณะย่อย ๆ หรือลักษณะอื่นร่วมอยู่ด้วย ถ้านำแบบทดสอบปลายเปิด ที่มี 4 องค์ประกอบตามที่ได้จากการวิเคราะห์องค์ประกอบ ไปออกข้อสอบเป็นแบบเลือกตอบแล้วตัดให้เหลือ 4 องค์ประกอบเท่ากัน แล้วกำหนดลักษณะของแบบทดสอบชุดนี้ใหม่ ตามองค์ประกอบที่ได้ ซึ่งจะทำให้ได้ค่าความเที่ยงตรงเชิงเหมือนและเชิงจำแนกเด่นชัดขึ้น (Marsh. : 1983) อย่างไรก็ตาม หลักฐานความเที่ยงตรงวิธีนี้ เป็นการคำนวณจากคะแนนสอบ และยังเป็นค่าความเที่ยงตรงร่วมกับเครื่องมือวัดอื่น คือแบบทดสอบเลือกตอบ ดังนั้นจึงต้องพิจารณาเป็นเพียงหลักฐานหนึ่งเท่านั้น และควรพิจารณาร่วมกับวิธีอื่นด้วย

3. ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร ρ_{TS} ของไฮเซนและบรอนสเตอร์ (Heise and Bohmstedt . 1970) มีค่าเท่ากับ .8993 แสดงว่าแบบทดสอบปลายเปิดวิชาคณิตศาสตร์มี ความเที่ยงตรงเชิงโครงสร้างสูง นั่นคือ สามารถวัดได้ตรงตามที่ต้องการวัด และสามารถนำไปใช้ประเมินผลความสามารถของผู้เข้าสอบได้ และเป็นค่าความเที่ยงตรงที่น่าเชื่อถือได้มาก เพราะการวิเคราะห์นั้น ได้คำนวณหาความสัมพันธ์จากคะแนนจริง กับคะแนนน้ำหนักความสำคัญขององค์ประกอบ และวิธีดังกล่าวได้มี ผู้นำไปใช้ในการแสดงค่าความเที่ยงตรงเชิงโครงสร้าง ในกรณีที่เป็นแบบทดสอบแยกวัดเป็นองค์ประกอบ เช่น สุทธิพงศ์ บุญผดุง(2541 : 72) ที่สร้างแบบทดสอบวัดลักษณะความรับผิดชอบประกอบด้วยคุณลักษณะ 4 ด้าน คือ ด้านความมีวินัยในตนเอง ด้านความเชื่อมั่นในตนเอง ด้านความกระตือรือร้น และด้านความขยันหมั่นเพียร อดทน อดกลั้น พบว่า ความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบที่คำนวณด้วยสูตร ρ_{TS} ของไฮเซนและบรอนสเตอร์ (Heise and Bohmstedt . 1970) มีค่าเท่ากับ .87864 ซึ่งค่าที่ได้อยู่ในระดับที่สูง

4. ความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตร Ω_w ของอัลเลน (Allen. 1974) มีค่าเท่ากับ .8755 และที่คำนวณด้วยสูตร r_{12} ของเลียว (Liou.1989) มีค่าเท่ากับ .6580 ซึ่งจะเห็นได้ว่าค่าแตกต่างกันมาก ทั้งนี้อาจเป็นเพราะว่าแบบทดสอบปลายเปิดชุดนี้มุ่งวัดองค์ประกอบ 3 ด้านแยกกันคือ ความเข้าใจในความคิดรวบยอด การแก้ปัญหา และการให้เหตุผลทางคณิตศาสตร์ ซึ่งสูตรโอเมก้าที่ถ่วงน้ำหนักของอัลเลน (Allen) นั้น ได้พัฒนาขึ้นมาโดยอาศัยข้อตกลงดังกล่าว ส่วนสูตรของเลียวนั้น พิจารณาจากข้อตกลงที่ว่าแบบทดสอบวัดองค์ประกอบเดียวกันร่วมกัน ดังนั้น สูตรของอัลเลน (Allen) จึงให้ค่าความเชื่อมั่นสูงกว่า จากผลดังกล่าว ยังแสดงถึงความเที่ยงตรงเชิงโครงสร้างได้อีกทางหนึ่งด้วย ดังนั้น แบบทดสอบที่ประกอบด้วยหลายองค์ประกอบ จึงควรคำนวณด้วยสูตรของอัลเลน และวิธีดังกล่าวได้มี ผู้นำไปใช้ในการแสดงค่าความเชื่อมั่นของแบบทดสอบ ในกรณีที่แบบทดสอบ แยกวัดเป็นองค์ประกอบ เช่น งานวิจัยของวรพร สกลจิตรานนท์ ที่วิจัยเกี่ยวกับการประมาณค่าความเชื่อมั่น ของแบบทดสอบความถนัดทางการเรียน ด้วยสูตร Ω , Ω_w , θ และ θ_k โดยแบ่งแบบทดสอบความถนัดทางการเรียน เป็น 8 ด้าน แล้วคำนวณค่าความเชื่อมั่น ด้วยสูตรดังกล่าว ผลปรากฏว่า ค่าความเชื่อมั่นของแบบทดสอบความถนัดทางการเรียน ที่แบ่ง

เป็น 8 ด้าน ที่คำนวณด้วยสูตร Ω_w มีค่าสูงสุด แสดงให้เห็นว่าแบบทดสอบที่ประกอบด้วยหลายองค์ประกอบ
ควรหาค่าความเชื่อมั่นที่คำนวณด้วยสูตรของอัลเลน

ข้อเสนอแนะ

ผลการวิจัยครั้งนี้ สามารถสรุปเป็นข้อเสนอแนะในการนำไปใช้ และในการวิจัย ได้ดังนี้

1. ข้อเสนอแนะในการนำไปใช้

1.1 แบบทดสอบที่ประกอบด้วยหลายองค์ประกอบ ควรหาค่าความเชื่อมั่นที่คำนวณด้วยสูตร
 Ω_w ของอัลเลน

1.2 ควรใช้แบบทดสอบปลายเปิดในวิชาคณิตศาสตร์ เพราะสามารถวัดความคิดของผู้สอบได้
เที่ยงตรงกว่าแบบทดสอบเลือกตอบ

2. ข้อเสนอแนะในการวิจัย

2.1 ควรทำการศึกษาด้านเจตคติ ควบคู่กับการใช้แบบทดสอบปลายเปิด

2.2 ควรทำการศึกษาเปรียบเทียบการตรวจให้คะแนนแบบแยกองค์ประกอบ และแบบการ
ประเมินรวม ว่าวิธีใดจะให้ค่าความเชื่อมั่นสูงกว่ากัน

2.3 ควรทำการศึกษาในกรณีที่ให้ผู้ตรวจมากกว่า 1 คน จะส่งผลต่อความเที่ยงตรงและ
ความเชื่อมั่น อย่างไร

2.4 ควรศึกษาเปรียบเทียบเกี่ยวกับค่าใช้จ่าย และเวลาในการสอบด้วยแบบทดสอบปลายเปิด

บรรณานุกรม

บรรณานุกรม

- โกวิท ประวาลพุกษ์. การพัฒนาทรัพยากรมนุษย์สำหรับอนาคต. กรุงเทพฯ : โรงพิมพ์การศาสนา, 2534.
- คณะกรรมการการประถมศึกษาแห่งชาติ, สำนักงาน. การวัดและการประเมินผลของนักเรียน : เอกสารประกอบชุดการสอนเพื่อการเรียนรู้ที่แท้จริง. กรุงเทพฯ : สำนักงาน คณะกรรมการการประถมศึกษาแห่งชาติ, 2539.
- เจริญ คำยัง. "การประเมินสภาพจริง (Authentic Assessment)," เอกสารการสัมมนาเชิงปฏิบัติการ ครูคณิตศาสตร์ กลุ่ม 4 กรุงเทพฯ : 28-29 เมษายน 2540. สำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ.
- ชวาล แพทย์กุล. เทคนิคการวัดผล. กรุงเทพฯ : ไทยวัฒนาพานิช, 2516.
- _____. เทคนิคการเขียนคำถามเลือกตอบ. กรุงเทพฯ : โรงพิมพ์เจริญผล, 2530.
- ธนารักษ์ ปิ่นเทียน. "การวัดประเมินผลสำหรับการเรียนการสอนแนวใหม่," ครูเชียงราย. 33(175) : 27-29 ; มกราคม 2540.
- _____. "การประเมินผลจากสภาพจริง," ครูเชียงราย. 33(181) : 4-8 ; กรกฎาคม 2540.
- บรรทม มณีโชติ. การศึกษารูปแบบของข้อคำถามวัดลักษณะนิสัยด้านความเสียสละชนิดข้อความและชนิดสถานการณ์ ที่มีต่อคุณภาพของแบบทดสอบ. ปรียุญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2530. อัดสำเนา.
- บุญชม ศรีสะอาด. "วิเคราะห์อำนาจจำแนก," การวัดผลการศึกษา. 11(31) : 39-48 ; พฤษภาคม-สิงหาคม 2532.
- บุญเชิด ภิญโญนนตพงษ์. การสร้างแบบทดสอบ. ภาควิชาพื้นฐานการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2525.
- _____. "Congeneric Part Reliability," การวัดผลการศึกษา. 12(34) : 28-31 ; พฤษภาคม-สิงหาคม 2533.
- _____. "Congeneric Model," การวัดผลการศึกษา. 11(42) : 8-20 ; มกราคม-เมษายน 2536.
- _____. การประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งส่วนย่อยตามแบบจำลองคะแนนจริงล้มพันธ์. ปรียุญานิพนธ์ กศ.ด. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2537. อัดสำเนา.
- _____. "การประเมินผลแนวใหม่ : พอทโฟลิโอ," ศึกษาศาสตร์. 1(1) : 59-67 ; กันยายน 2538 .
- เบญจวรรณ แยมละมูล. การศึกษาความเที่ยงตรงแบบหลายลักษณะ-หลากวิธี ของแบบทดสอบวัดเจตคติที่มีต่อโรงเรียนของนักเรียนและครูในระดับประถมศึกษา. ปรียุญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2538. อัดสำเนา.
- ปราโมทย์ มากชู. "ยุทธวิธีการฝึกแก้โจทย์ปัญหาคณิตศาสตร์ ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 ในเขตภาคเหนือตอนล่าง," มหาวิทยาลัยนเรศวร. 5(1) : 28-39 ; มกราคม-มิถุนายน 2540.
- พรรณี เทพสุทร. การสร้างแบบทดสอบวัดความวิตกกังวลในการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1. ปรียุญานิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2537. อัดสำเนา.

- พวงรัตน์ ทวีรัตน์. การสร้างและพัฒนาแบบทดสอบวัดผลสัมฤทธิ์. กรุงเทพฯ : สำนักทดสอบทางการศึกษา และจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ, 2531.
- ไพศาล หวังพานิช. การวัดผลการศึกษา. สำนักทดสอบ มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2525.
- ภัทรกุล จิรวิทย์านนท์. "ผลสัมฤทธิ์วิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ของไทยเทียบกับนานาชาติ," สสวท. 15(2) : 11-14 ; เมษายน-มิถุนายน 2530.
- ยุพิน พิพิธกุล. การนิเทศการสอนวิชาคณิตศาสตร์. กรุงเทพฯ : คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2537.
- วิชาการ, กรม, กระทรวงศึกษาธิการ. รายงานผลการวิจัยข้อบกพร่องในการเรียนวิชาคณิตศาสตร์. กรุงเทพฯ : โรงพิมพ์การศาสนา, 2532.
- _____. หลักสูตรมัธยมศึกษาตอนต้น พุทธศักราช 2521(ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ : โรงพิมพ์การศาสนา, 2535.
- _____. รายงานผลการประเมินคุณภาพการศึกษาปีการศึกษา 2538. กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว, 2538.
- _____. การประเมินผลจากสภาพจริง. กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว, 2539.
- วรพร สกุลจิตรานนท์. การประมาณค่าความเชื่อมั่นของแบบทดสอบความถนัดทางการเรียนด้วยสูตร Ω สูตร Ω_w สูตร θ และสูตร θ_k . ปรินิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2541. อัดสำเนา.
- ส.วาสนา ประवालพฤษ. "การประยุกต์ใช้การวัดและการประเมินความสามารถจริงในสภาพการเรียนการสอน," การวัดผลการศึกษา. 17(51) : 32-41 ; มกราคม-เมษายน 2539.
- _____. "การวัดผลจากการปฏิบัติจริง," สารพัฒนหลักสูตร. 15(125) : 46-50 ; เมษายน-มิถุนายน 2539.
- สมสว่าง ธนะพานิชยกุล. การสร้างแบบทดสอบวัดกระบวนการแก้ปัญหาโจทย์คณิตศาสตร์. ปรินิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2539. อัดสำเนา.
- เสกสิทธิ์ แสนทวีสุข. การศึกษาความเที่ยงตรงเชิงโครงสร้างของแบบวัดมโนภาพแห่งตนหลายมิติ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3. ปรินิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2539. อัดสำเนา.
- สุทธิพงศ์ บุญผดุง. การสร้างแบบทดสอบวัดลักษณะความรับผิดชอบ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 โดยใช้การแสดงหลักฐานความเที่ยงตรง ความไม่เที่ยงตรง และความเชื่อมั่น. ปรินิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2541. อัดสำเนา.
- สุภาวดี ตั้งบุผา. การสร้างแบบทดสอบวัดความสามารถในการคิดสร้างสรรค์ทางคณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 กรุงเทพมหานคร. ปรินิพนธ์ กศ.ม. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2533. อัดสำเนา.
- สำเร็จ บุญเรืองรัตน์. ลักษณะหลากหลายวิธีหลาย : การประยุกต์ใช้วิเคราะห์ความเที่ยงตรงของแบบทดสอบ. กรุงเทพฯ : สยามศึกษา, 2528.
- อนันต์ ศรีโสภณ. การวัดผลการศึกษา. ไทยวัฒนาพานิช, 2524.

- Aiken, Lewis R. "Testing with Multiple - Choice Item," Journal of Research and Development in Education. 20(4) : 45-58 ; 1987.
- Allen, M. J., and Yen W.M. Introduction to Measurement Theory. Monterey : Brook/Cole Publishing , 1979.
- Allen, M.P. "Construction of composite measures by the canonical factor regression method," Sociological Methodology. San Francisco : Jossey Bass, 1974.
- Arlin, P.K. "A Multitrait-Multimethod Validity Study of a Test of Formal Reasoning," Educational and Psychological Measurement. 42(4) : 1077-1088 ; 1982.
- Bacon, D.R., Saure P.L. and Yong M. "Composite Reliability in Structural Equation Modeling," Educational and Psychological Measurement. 55(3) : 394 - 405 ; 1995.
- Barbour, J.D. " A Multitrait-Multimethod Study of the Construct Validity of Frequently Use Behavior Rating Scales for Hyperactivity," Dissertation Abstracts International. 50(7) : 1981 ; January, 1990.
- Birenbaum M. and Tasuoka K.K. "Open-ended Versus Multiple-Choice Response Formats-It Does Make a Difference for Diagnostic Purposes," Applied Psychological Measurement. 11(4) : 385-395 ; December, 1987.
- Bucklew, J.E. "Convergent and Discriminant Validation of Self-Report and Behavioral Rating Measures of Depression and Aggression Symptoms For Children in Residential Treatment," Dissertation Abstracts International. 49(1) : 234 ; July, 1988.
- Centra, John A. "Validation by the Multigroup-Multiscale Matrix : An Adaptation of Campbell and Fiske's Convergent and Discriminant Validation Procedure," Educational and Psychological Measurement. 31(3) : 675-683 ; 1971.
- Crocker, Linda and Algina, Jame. Introduction to Classical and Modern Test Theory. New York : Holt, Rinehart and winston, 1986.
- Gronlund, Norman E. Measurement and Evaluation in Teaching. New York : Macmillan, 1965.
- Heise,D.R. and G.W. Bohrnstedt. "Validity Invalidity and Reliability," Sociological Methodology. San Francisco : Jossey Bass, 1970.
- Khan, S.B. "A Comparative Study of Assessing Children's School Related Attitudes," Journal of Educational Measurement. 1 : 59-66 ; 1978.
- Kristof, W. "Estimation of Reliability and True Score Variance from a Split of a Test into Three Arbitrary Part," Psychometrika. 39(4) : 491-499 ; 1974.
- Linn, R.L., Baker, E. and Dunbar, S.W. "Complex Performance based assessment : Expectations and Validation criteria, Educational Research. 20 : 15-21 ; 1991.
- Liou, M. "A Anote on Reliability Estimation for a test with Component of Unknown Functional Length," Psychometrika. 54 : 153-163 ; 1989.
- Magone, E. "Validating the Cognitive Complexity and Content Quality of a Mathematics Performance Assessment," International Journal of Education Research. 21(3) : 317-331 ; 1994.

- Marsh, H.W. and Smith, I.D. "Multitrait-Multimethod Analyses of Two Self-Concept Instruments," Journal of Educational Psychology. 74(3) : 430-440 ; 1982.
- Marsh, H.W., Smith, I.D. and Barnes, J. "Multitrait-Multimethod Analysis of the Self-Description Questionnaire : Student-Teacher agreement on Multidimensional Rating of Student Self-Concept," American Educational Research Journal. 20(3) : 333-357 ; 1983.
- Marshall, J.E. "An Investigation of the Construct Validity of the Test of Basic Skills in Science : A Multitrait-Multimethod Analysis (Science Process Skills)," Dissertation Abstracts International. 51(11) : 36-91; May, 1991.
- National Council of Teachers of Mathematics. Curriculum and evaluation standards for school mathematics. Reston, VA : The council. 1989.
- Nitko , Anthong J. Educational Assessment of Students. 2 nd ed. Englewood Cliffs, N.J. : Merrill, 1996.
- Noll , Victor H. Introduction to Education Measurement. Boston : Houghton, Mifflin, 1957.
- Powell , J.C., "The Interpretation of Wrong Answer form a Multiple - Choice Test," Education and Psychological Measurement. 28(2) : 403-412 ; Summer, 1968.
- Reys, Robert E. Helping Children Learn Mathematics. 3rd ed. Boston : Allyn and Bacon, 1992.
- Singer, J. "Measurement of Self-Concept Among Disabled Children : A Construct Validation Study of Self-Description Questionnaire," Dissertation Abstracts International. 49(10) : 3006-A ; April, 1989.
- Thorndike. R.L. Education Measurement. Washington DC : American Council on Education, 1971.
- Trant, T.J. "Convergent and Discriminant Validity of the Leadership Skill Inventory," Dissertation Abstracts International. 47(11) : 4034 ; May, 1987.
- Tryjankowski, E.M. "Convergent - Discriminant Validity of the Jewish Evaluation and Vocational System," Dissertation Abstracts International. 17(2) : 516 - A ; August, 1986.
- Wainwright, O.D. "Measurement of Self - Concept Among Disabled Children : A Construct Validation Study of the Self - Description Questionnaire," Dissertation Abstracts International. 49(10) : 3006-A ; April, 1989.

ภาคผนวก

รายชื่อผู้เชี่ยวชาญ

- | | |
|---|--|
| 1. รองศาสตราจารย์นิภา ศรีไพโรจน์ | ภาควิชาการวัดผลและวิจัยการศึกษา
คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ประสานมิตร |
| 2. อาจารย์ชวลิต รวยอาจิม | ภาควิชาการวัดผลและวิจัยการศึกษา
คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ประสานมิตร |
| 3. ดร. ณัฏฐวรรณ เสวตมัลย์ | ภาควิชาหลักสูตรและการสอน คณะศึกษาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร |
| 4. รองศาสตราจารย์ ดร. พิชากร แปลงประสพโชค | ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร |
| 5. อาจารย์ชนันท์ตา นัทรทอง | โรงเรียนบดินทร์เดชา (สิงห์ สิงหเสนี) |

เกณฑ์การให้คะแนนแบบทดสอบปลายเปิดวิชาคณิตศาสตร์

ข้อ	พฤติกรรมที่ต้อง การวัด	ข้อสอบ	คะแนน
1	ความคิดรวบยอด เกี่ยวกับจำนวน เต็ม	ถ้า x เป็นจำนวนใด ๆ แล้ว $\frac{-x}{3}$ จะเป็นจำนวนเต็มแต่ละประเภทเมื่อ x มีค่าเท่าใด	0- ไม่ตอบหรือตอบไม่ถูกต้อง 1- แสดงวิธีคิดเล็กน้อยแต่ยังไม่ได้คำตอบ เช่นบอกได้ว่า x เป็นจำนวนเต็ม 2- ยกตัวอย่างคำตอบถูกต้อง อธิบาย/แปล ผลถูกต้อง 1 คำ ตอบ 3- คำตอบถูกต้อง อธิบาย/แปลผลถูกต้อง 2 คำตอบ 4- คำตอบถูกต้องสมบูรณ์ อธิบาย/แปลผลถูกต้อง ทุกขั้นตอน
2	ความเข้าใจในกฎ หรือหลักการทาง คณิตศาสตร์	เมื่อ A และ B เป็นจำนวน และ $*$ เป็น เครื่องหมายที่ใช้ทำนองเดียวกับเครื่องหมาย $+, -, \times, \div$ ถ้ากำหนดค่าให้ $A*B = A - (B \div A)$ แล้ว $(0.16)*(-2)$ มีค่าเท่าใด	0- ไม่ตอบหรือแทนค่าไม่ถูกต้อง และไม่ทำอะไรต่อ 1- แทนค่าถูกเกือบสมบูรณ์ 2- แทนค่าถูกต้องคำนวณผิดพลาด มาก 3- แทนค่าถูกต้อง คำนวณผิดพลาด เล็กน้อย 4- คำตอบถูกต้องสมบูรณ์ แสดงแนว คิดชัดเจน
3	ความคิดรวบยอด เกี่ยวกับการคูณ จำนวนเต็ม	ถ้า p เป็นจำนวนเต็มบวก q เป็น จำนวนเต็มลบ และ $p \times q \times r$ เป็น จำนวนเต็มบวก แล้ว $2r$ เป็น จำนวน เต็มประเภทใด เพราะเหตุใด	0- ไม่ตอบหรือตอบไม่ถูกต้อง 1- แสดงวิธีคิดเล็กน้อย แต่มีแนวโน้ม ไปสู่คำตอบที่ถูกต้อง 2- ตอบได้ว่า r เป็นจำนวนเต็ม ประเภทใดเพราะอะไร 3- คำตอบถูกต้อง เหตุผลไม่สมบูรณ์ 4- คำตอบถูกต้อง เหตุผลถูกต้องครบ ทุกขั้นตอน

เกณฑ์การให้คะแนน(ต่อ)

ข้อ	พฤติกรรมที่ต้อง การวัด	ข้อสอบ	คะแนน
4	การแก้ปัญหา จำนวนเต็ม	สี่เหลี่ยมผืนผ้ารูปหนึ่งมีพื้นที่ 24 ตาราง- เซนติเมตร และความยาวแต่ละด้านเป็น จำนวนนับที่ 2หารลงตัวจงหาความยาว ด้านของรูปสี่เหลี่ยมผืนผ้าที่เป็นไปได้ ทั้งหมด	0- ไม่ตอบ หรือตอบไม่ถูกต้องและไม่ แสดงการคิดใด ๆ 1- เขียนสูตรการหาพื้นที่สี่เหลี่ยมผืน ผ้าได้ถูกต้อง แต่แทนค่าไม่ถูกต้อง 2- แทนค่าในสูตรได้ถูกต้อง แต่ไม่ได้ ตอบคำถาม 3- แทนค่าในสูตรได้ถูกต้อง ตอบคำ ถามได้ไม่ครบทุกกรณี 4- แทนค่าในสูตรได้ถูกต้อง ตอบคำ ถามได้ครบทุกกรณี
5	การแก้ปัญหา ทศนิยม	แป้งข้าวเหนียวที่บรรจุถุงละ 2 กิโลกรัม ราคาถุงละ 15.50 บาท และบรรจุถุงละ 5 กิโลกรัม ราคาถุงละ 37.75 บาท ถ้า ต้องการแป้งข้าวเหนียวทั้ง 2 ขนาด จำนวน 19 กิโลกรัม โดยจ่ายเงินน้อยที่สุด จะได้ยี่ห้อใด	0- ไม่ตอบ,ตอบไม่ถูกต้องและไม่ แสดงการคิดใด ๆ 1- คัดลอกข้อมูลที่บอกถึงความเข้าใจใน ปัญหาได้เล็กน้อยแต่ไม่สามารถ ดำเนินการแก้ปัญหาได้ หรือ พยายามที่จะใช้ข้อมูลในปัญหามา ใช้ในการแก้ปัญหาแม้ว่าแนวทาง ที่ว่าจะไม่นำ ไปสู่คำตอบที่ถูกต้อง ต้อง 2- หาคำตอบได้ 1 กรณีแสดงการคิด คำนวณชัดเจน 3- แสดงการแก้ปัญหาได้มากกว่า 1 กรณี แต่คำนวณผิดพลาดเล็กน้อย 4- แสดงการแก้ปัญหาได้ถูกต้อง สมบูรณ์ มากกว่า 1 กรณี และ เลือกคำตอบที่ถูกต้องได้ถูกต้อง

เกณฑ์การให้คะแนน(ต่อ)

ข้อ	พฤติกรรมที่ต้อง การวัด	ข้อสอบ	คะแนน				
6	แก้ปัญหา ทศนิยม	อัตราค่าโดยสารรถประจำทางเป็นดังนี้ <table border="1"><tr><td>ไป/กลับอย่างเดียว</td><td>3.50 บาท</td></tr><tr><td>ตั๋วเดือน(ไป-กลับ/เที่ยว/วัน)</td><td>184.50 บาท</td></tr></table> นุก เรียนอยู่โรงเรียนแห่งหนึ่งในเดือน กันยายน นุกจะไปโรงเรียนด้วยรถประจำ ทางทั้งไป-กลับ ในวันจันทร์ วันพุธและวัน ศุกร์ ส่วนวันอังคารและวันพฤหัสบดี นุกจะ ไปโรงเรียนด้วยรถประจำทางเฉพาะตอนเช้า ตอนเย็นคุณแม่ไม่รับ ถ้าวันที่ 1 กันยายน ตรงกับวันอังคาร และไม่มีวันสำคัญใด ๆ นัก เรียนคิดว่า นุกควรจะซื้อตั๋วเดือนหรือไม่ เพราะเหตุใด	ไป/กลับอย่างเดียว	3.50 บาท	ตั๋วเดือน(ไป-กลับ/เที่ยว/วัน)	184.50 บาท	0- ไม่ตอบ,คัดลอกข้อมูลที่ไม่สำคัญและไม่ได้ นำมาใช้ในการแก้ปัญหา 1- แสดงวิธีคิดเล็กน้อย แต่ มีแนวโน้มที่จะ นำไปสู่คำตอบที่ถูกต้อง 2- ตอบว่าควรซื้อตั๋วเดือน/ไม่ควรซื้อ อย่างไร อย่างหนึ่ง เนื่องจากจำนวนผิดพลาด มาก หรืออธิบายไม่สอดคล้องกับหลัก ความเป็นจริง 3- ไม่ควรซื้อตั๋วเดือน จำนวนผิดพลาดเล็ก น้อย และอธิบายได้สอดคล้องกับหลัก ความเป็นจริง 4- ไม่ควรซื้อ หรือตอบว่าเป็นไปได้ทั้ง 2 กรณี และในแต่ละ กรณีสามารถอธิบาย ได้ถูกต้องชัดเจนและ สอดคล้องกับหลัก ความเป็นจริง
ไป/กลับอย่างเดียว	3.50 บาท						
ตั๋วเดือน(ไป-กลับ/เที่ยว/วัน)	184.50 บาท						

เกณฑ์การให้คะแนน(ต่อ)

ข้อ	พฤติกรรมที่ต้อง การวัด	ข้อสอบ	คะแนน
7	การแก้ปัญหา เศษส่วน	ในการเลือกตั้งหัวหน้าห้องของนักเรียนห้องหนึ่ง มีผู้สมัครรับเลือกตั้ง 2 คน ปรากฏว่า $\frac{1}{5}$ ของจำนวนบัตรทั้งหมดเป็นบัตรเสีย $\frac{3}{4}$ ของบัตรที่ใช้ได้ลงคะแนนให้ผู้ชนะ และ ผู้ชนะได้คะแนนมากกว่าผู้แพ้ 18 คะแนน จงหาว่ามีนักเรียนมาใช้สิทธิ์ลงคะแนนเสียงทั้งหมดกี่คน	0- ไม่ตอบ หรือตอบไม่ถูกต้องและไม่แสดงการคิดเลย หรือลอกข้อมูลในปัญหาแต่ไม่ได้นำมาใช้ในการแก้ปัญหา 1- คัดลอกข้อมูลที่บอกถึงความเข้าใจในปัญหาได้เล็กน้อยแต่ไม่สามารถดำเนินการแก้ปัญหาได้ 2- แก้ปัญหาส่วนย่อยได้บ้าง แต่ไม่สามารถนำไปสัมพันธ์กันได้ หรือสามารถเชื่อมโยงรายละเอียดที่สำคัญของปัญหาได้บ้างแต่คำนวณผิดพลาดมาก 3- แก้ปัญหาส่วนย่อยและเชื่อมโยงปัญหาได้เกือบสมบูรณ์ 4- คำตอบถูกต้อง เลือกยุทธวิธีเหมาะสม กระบวนการชัดเจน หรือใช้วิธีที่แปลกใหม่แต่คำนวณผิดพลาดเล็กน้อย
8	การให้เหตุผล ในรูปของการหา ความสัมพันธ์	จงเติมจำนวนในช่องว่างต่อจากจำนวนที่กำหนดให้ให้ถูกต้อง และจำนวนแต่ละจำนวนสัมพันธ์กันอย่างไร $x-4$, $x-2$, , $x+2$, $x+4$ จำนวนใน คือ แสดงวิธีคิด สรุปความสัมพันธ์ 	0- ไม่ตอบ, ตอบไม่ถูกต้องและไม่แสดงการคิดใดๆ 1- คำตอบไม่ถูกต้อง แสดงวิธีคิดเล็กน้อย ไม่สามารถสรุปความสัมพันธ์ได้ 2- แสดงแนวคิดได้บ้าง แต่มีแนวโน้มนำไปสู่คำตอบที่ถูกต้องและสรุปความสัมพันธ์ได้ 3- คำตอบถูกต้อง วิธีคิด และสรุปความสัมพันธ์ได้ไม่สมบูรณ์ 4- คำตอบถูกต้อง วิธีคิด และสรุปความสัมพันธ์ได้ถูกต้อง ใช้ภาษาได้ชัดเจน

ประวัติย่อผู้วิจัย

ชื่อ-ชื่อสกุล	นางสาวอรอนงค์ บำรุง
วัน เดือน ปีเกิด	23 มีนาคม 2505
สถานที่อยู่ปัจจุบัน	35/186 ลาดพร้าววิลเลจ ซอยลาดพร้าว 71 แขวงลาดพร้าว เขตลาดพร้าว กรุงเทพมหานคร
สถานที่ทำงาน	โรงเรียนพระโขนงพิทยาลัย แขวงบางจาก เขตพระโขนง กรุงเทพมหานคร
ตำแหน่ง	อาจารย์ 2 ระดับ 6

ประวัติการศึกษา

พ.ศ. 2523	มัธยมศึกษาตอนปลาย โรงเรียนสายน้ำผึ้ง กรุงเทพมหานคร
พ.ศ. 2527	คบ. (เอกคณิตศาสตร์) วิทยาลัยครูสวนสุนันทา
พ.ศ. 2542	กศ.ม.(การวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร

การแสดงผลหลักฐานความเที่ยงตรงและความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์

บทคัดย่อ
ของ
อรอนงค์ บำรุง

เสนอต่อมหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
พฤษภาคม 2542

การศึกษานี้เพื่อ แสดงหลักฐานความเที่ยงตรง และความเชื่อมั่นของแบบทดสอบปลายเปิด วิชาคณิตศาสตร์ โดยมีจุดมุ่งหมายเพื่อ 1) แสดงหลักฐานความเที่ยงตรงเชิงโครงสร้าง ของแบบทดสอบ ปลายเปิดวิชาคณิตศาสตร์ 2 วิธี คือ หาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) และ หาความสัมพันธ์ระหว่างคะแนนจริงกับคะแนนองค์ประกอบ ที่คำนวณด้วยสูตรของ ไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt) และ 2) หาค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ 2 วิธี คือ วิเคราะห์น้ำหนักองค์ประกอบ ที่คำนวณด้วยสูตรของอัลเลน(Allen) และวิเคราะห์แบบคะแนนจริงสัมพันธ์ ที่คำนวณด้วยสูตรของเลียว(Liou)

กลุ่มตัวอย่างที่ใช้ในการศึกษานี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2541 ของโรงเรียนสังกัดกรมสามัญศึกษา กลุ่ม 4 กรุงเทพมหานคร จำนวน 300 คน โดยใช้การสุ่มแบบ 2 ขั้นตอน

ผลการวิจัย พบว่า

1. หลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่หาด้วยวิธี หาความสัมพันธ์หลายลักษณะ-หลายวิธี (Multitrait - Multimethod) พบว่า สัมประสิทธิ์สหสัมพันธ์ของ ความเที่ยงตรงเชิงเหมือน มีค่าตั้งแต่ .4862 ถึง .7316 และสัมประสิทธิ์สหสัมพันธ์ของความเที่ยงตรง เชิงข้ามแนก มีค่าตั้งแต่ .1508 ถึง .4753
2. หลักฐานความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วย สูตรของ ไฮเซนและบอร์นสเตดท์ (Heise and Bohmstedt) มีค่าเท่ากับ .8993
3. ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของอัลเลน(Allen) มีค่าเท่ากับ .8755
4. ค่าความเชื่อมั่นของแบบทดสอบปลายเปิดวิชาคณิตศาสตร์ ที่คำนวณด้วยสูตรของเลียว(Liou) มีค่าเท่ากับ .6580

VALIDITY EVIDENCE AND RELIABILITY OF OPEN-ENDED QUESTIONS IN MATHEMATICS

**AN ABSTRACT
BY
ORNANONG BUMROONG**

**Presented in partial fulfillment of the requirement for
the Master of Education degree in Educational Measurement
at Srinakharinwirot University
May 1999**

The research of this study was to investigate the validity evidence and the reliability of the open-ended questions in mathematics. The major purpose of this study were 1) to show the evidence of construct validity by using : Multitrait - Multimethod, and analysis method of Heise and Bohmstedt. 2) to investigate the test reliability by using : Weighted - Omega formula, and r_{L2} coefficient (Liou).

The sample of this study was 300 Mathayom Suksa II students in the secondary school under the Department of General Education in the fourth group of Bangkok, in the first semester in 1998 and was selected by two stage cluster sampling.

The findings were as followings:

1. The construct validity of the open-ended questions in mathematics by using Multitrait-Multimethod show that the correlation coefficient of convergent validity was ranged from .4862 to .5361 and the correlation coefficient of divergent validity was ranged from .1508 to .4753.
2. The construct validity of the open-ended questions in mathematics calculated by the formula of Heise and Bohmstedt was .8993.
3. The reliability of the open-ended questions in mathematics calculated by Weighted - Omega formula was .8755
4. The reliability of the open-ended questions in mathematics calculated by r_{L2} coefficient (Liou) was .6580.