การเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนด้วยวิธีที่แตกต่างกัน

นางสาวทัตมณี ชูขวัญ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2548
ISBN 974-14-3746-3
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A COMPARISON OF THE QUALITY OF MATHEMATICS CONCEPTUAL KNOWLEDGE ASSESSMENTS USING CONCEPT MAPPING WITH DIFFERENT SCORING METHODS

Miss Tadmanee Chukhuan

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education Program in Educational Measurement and Evaluation
Department of Educational Research and Psychology
Faculty of Education
Chulalongkorn University
Academic Year 2005

ISBN 974-14-3746-3

หัวข้อวิทยานิพนธ์	การเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนด้วยวิธีที่แตกต่างกัน
โดย	นางสาวทัตมณี ชูขวัญ
สาขาวิชา	การวัดและประเมินผลการศึกษา
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

..... คณบดีคณะครุศาสตร์
(รองศาสตราจารย์ ดร.พฤทธิ์ ศิริบรรณพิทักษ์)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

..... อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์)

..... กรรมการ
(รองศาสตราจารย์ ดร.ศิริเดช สุชีวะ)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ทัตมณี ชูขวัญ : การเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนด้วยวิธีที่แตกต่างกัน. (A COMPARISON OF THE QUALITY OF MATHEMATICS CONCEPTUAL KNOWLEDGE ASSESSMENTS USING CONCEPT MAPPING WITH DIFFERENT SCORING METHODS) อ.ที่ปรึกษา : ผศ. ดร.เอมอร จังศิริพรปกรณ์, 157 หน้า. ISBN 974-14-3746-3.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) ศึกษาและพัฒนาเกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ด้วยวิธีของโนแวคและโกวิน และวิธีของบอลท์ 2) เพื่อเปรียบเทียบคุณภาพของการประเมินเมื่อตรวจด้วยวิธีที่แตกต่างกัน โดยพิจารณาจาก ความตรงตามสภาพ ความตรงเชิงโครงสร้าง ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ ความเที่ยงที่แทนด้วยสัมประสิทธิ์การสรุปอ้างอิง และ 3) เพื่อวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสม กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 โรงเรียนสุรนารีวิทยา สำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต 1 จำนวน 220 คน และผู้ประเมินที่ตรวจให้คะแนนแผนผังมโนทัศน์ด้วยวิธีของโนแวคและโกวิน และวิธีของบอลท์ รวม 4 คน เครื่องมือที่ใช้ในการวิจัยคือ แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ และแบบทดสอบประเภทเลือกตอบ ข้อมูลที่ได้นำมาวิเคราะห์ค่าสถิติพื้นฐาน ตรวจสอบความตรงตามสภาพและความตรงเชิงโครงสร้างด้วยเมทริกซ์พหุลักษณะพหุวิธี โดยใช้โปรแกรม SPSS for Windows วิเคราะห์ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบด้วยโปรแกรม MULTILOG และวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงด้วยโปรแกรม GENOVA

ผลการวิจัยพบว่า

1. เกณฑ์การให้คะแนนของโนแวคและโกวิน พิจารณาจากแผนผังมโนทัศน์ในด้านคุณภาพ 4 ด้าน ซึ่งแต่ละด้านจะมีคะแนนเต็ม 4 คะแนน คือ 1.) ประพจน์ 2.) การจัดลำดับขั้นตอน 3.) การเชื่อมระหว่างมโนทัศน์ และ 4.) ตัวอย่าง และเกณฑ์การให้คะแนนของบอลท์ พิจารณาจากแผนผังมโนทัศน์ในด้านการจัดแผนผังมโนทัศน์ 6 คะแนน และความถูกต้อง 4 คะแนน

2. การเปรียบเทียบคุณภาพด้านความตรงและความเที่ยง

2.1 ค่าความตรงตามสภาพของคะแนนการประเมินแผนผังมโนทัศน์ เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินมีค่าสูงกว่าการตรวจให้คะแนนด้วยวิธีของบอลท์ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

2.2 แผนผังมโนทัศน์ เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน และการตรวจให้คะแนนด้วยวิธีของบอลท์ มีความตรงเชิงโครงสร้างทั้งสองวิธี

2.3 ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ ของวิธีการให้คะแนนของโนแวคและโกวินมีค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบโดยเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

2.4 ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์ เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินมีค่าสูงกว่าการตรวจให้คะแนนด้วยวิธีของบอลท์

3. จำนวนผู้ประเมินและจำนวนแผนผังที่เหมาะสม ที่ทำให้ค่าสัมประสิทธิ์สรุปอ้างอิงอยู่ในเกณฑ์เหมาะสมเมื่อกำหนดค่าสัมประสิทธิ์สรุปอ้างอิงเพื่อการนำไปใช้อิงเกณฑ์ เท่ากับ 0.8 พบว่าวิธีการให้คะแนนของโนแวคและโกวินต้องใช้ผู้ประเมินจำนวน 2 คน และจำนวนแผนผังมโนทัศน์จำนวน 8 แผนผัง หรือผู้ประเมิน 1 คน แผนผังมโนทัศน์ 12 แผนผัง และวิธีการให้คะแนนของบอลท์ต้องใช้ผู้ประเมิน 4 คน และจำนวนแผนผังมโนทัศน์ 12 แผนผัง หรือผู้ประเมิน 3 คน แผนผังมโนทัศน์ 16 แผนผัง ทั้งนี้ขึ้นอยู่กับผู้ปฏิบัติว่าจะเลือกใช้ให้เหมาะสมกับการเรียนการสอน

ภาควิชา วิจัยและจิตวิทยาการศึกษา ลายมือชื่อนิสิต .....

สาขาวิชา การวัดและประเมินผลการศึกษา ลายมือชื่ออาจารย์ที่ปรึกษา .....

ปีการศึกษา 2548

4783683027: MAJOR EDUCATIONAL MEASUREMENT AND EVALUTION

KEY WORD: SCORING METHOD/ CONCEPT MAPPING/ CONCURRENT VALIDITY/ CONSTRUCT VALIDITY/ ITEM INFORMATION FUNCTION/ TEST INFORMATION FUNCTION/GENERALIZABILITY COEFFICIENT.

TADMANEE CHUKHUAN : (A COMPARISON OF THE QUALITY OF MATHEMATICS CONCEPTUAL KNOWLEDGE ASSESSMENTS USING CONCEPT MAPPING WITH DIFFERENT SCORING METHODS. THESIS ADVISOR: ASSOC. PROF. AIMORN JANGSIRIPORNPAKORN, Ph.D., 157 pp. ISBN 974-14-3746-3.

The purposes of this research were 1) to study and develop scoring criterion of concept mapping assessment with the scoring method of Novak & Gowin and the scoring method of Bolte 2) to compare the quality of mathematics conceptual assessments using concept mapping with different scoring method ; concurrent validity, construct validity, item information function and test information function, the reliability from generalizability coefficient. 3) to study the optimal number of raters and number of concept mapping tasks when using the scoring method of Novak & Gowin or the scoring method of Bolte. The sample consisted of 220 Matayomsuksa 1 stududents of the Office of Nakhornratchasima Service Area Zone 1 and 4 raters who scored concept mapping tasks by using different scoring methods; Novak & Gowin scoring method and Bolte scoring method. The research instruments were concept mapping tasks and multiple choice test. SPSS for Windows was conducted to analyze basic statistics and MTMM matrix in order to find out the construct validity. The correlation between score form concept mapping and multiple choice test were analyzed to obtain the concurrent validity. MULTILOG was conducted to analyze GRM in order to find out item information function and test information function. GENOVA were analyzed to obtain the generalizability coefficient.

The findings were summarized as follows:

1. The scoring method of Novak & Gowin will based on 4 categories of concept mapping assessments which contain of 1.)proposition 2.) hierarchy 3.) cross link and 4.) examples (all contain 4 categories). The scoring method of Bolte will based on 6 and 4 categories of concept mapping assessments which contain of organization (6 categories) and accuracy (4 categories)

2. Comparison of validity and reliability

2.1 The scoring method of Novak & Gowin provided concurrent validity of concept mapping assessments larger than the scoring method of Bolte which were significantly different at .05 level.

2.2 Both the scoring method of Novak & Gowin and the scoring method of Bolte provided construct validity.

2.3 The scoring method of Novak & Gowin provided item information function and test in formation function of concept mapping assessments larger than the scoring method of Bolte which were significantly different at .05 level.

2.4 The scoring method of Novak & Gowin provided generalizability coefficient of concept mapping assessments larger than the scoring method of Bolte .

3. If Absolute Generalizability Coefficient was set at 0.8, the optimal number of raters and number of concept mapping tasks when using the scoring method of Novak & Gowin should be 2 raters with 8 concept mapping tasks or 1 rater with 12 concept mapping tasks. Meanwhile, the scoring method of Bolte should be 4 raters with 12 concept mapping tasks or 3 raters with 16 concept mapping tasks. Selection appropriate for instruction depends on user.

Department .....Educational Research and Psychology..... Student's signature.....Tadmanee Chukhuan

Field of study ....Educational Measurement and Evaluation. Advisor's signature.....Aimorn.....

Academic year .....2005.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สามารถเสร็จสมบูรณ์เป็นรูปเล่ม และเสร็จลุล่วงไปได้ด้วยความเมตตากรุณาจาก ผู้ช่วยศาสตราจารย์ ดร.เอมอร จังศิริพรปกรณ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ให้คำปรึกษาคำแนะนำ และแก้ไขปรับปรุงข้อบกพร่องต่าง ๆ ในการทำวิทยานิพนธ์เป็นอย่างดีมาโดยตลอด ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง ในความเมตตากรุณาของท่านไว้ ณ ที่นี้

ขอกราบพระคุณคณาจารย์ทุกท่าน ที่ได้ประสิทธิ์ประสาทความรู้อันเป็นประโยชน์กับผู้วิจัยในการทำวิทยานิพนธ์เป็นอย่างมาก ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงในความเมตตากรุณาของท่านไว้ ณ ที่นี้

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ ผู้ทรงคุณวุฒิทุกท่านที่ได้สละเวลาในการตรวจสอบเครื่องมือและขอขอบพระคุณผู้บริหารโรงเรียนและอาจารย์ผู้ปฏิบัติงานทุกท่านที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี

ขอขอบคุณ พี่ และเพื่อน ในภาควิจัยและจิตวิทยาการศึกษาทุกคนที่เป็นกำลังใจ และช่วยเหลือสนับสนุนการทำวิทยานิพนธ์ในด้านต่าง ๆ ต่อผู้วิจัยเสมอมา

สุดท้ายนี้ ผู้วิจัยขอกราบขอบพระคุณ คุณพ่อ คุณแม่ น้องชายและน้องสาวที่เป็นกำลังใจสำคัญ และส่งเสริมสนับสนุนด้านการเรียนของผู้วิจัยเสมอมาจนผู้วิจัยสำเร็จการศึกษา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย…… ง
บทคัดย่อภาษาอังกฤษ…….. จ
กิตติกรรมประกาศ……. ฉ
สารบัญ ……. ช
สารบัญตาราง..... ฌ
สารบัญภาพ..... ฎ
บทที่
1 บทนำ..... 1
ความเป็นมาและความสำคัญของปัญหา..... 1
คำถามการวิจัย..... 5
วัตถุประสงค์ของการวิจัย..... 5
ขอบเขตของการวิจัย..... 6
คำกำจัดความที่ใช้ในการวิจัย..... 7
ประโยชน์ที่คาดว่าจะได้รับ..... 8
2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... 10
ตอนที่ 1 ความคิดรวบยอดทางคณิตศาสตร์ 10
ตอนที่ 2 แผนผังมโนทัศน์..... 12
ตอนที่ 3 การตรวจสอบคุณภาพของเครื่องมือ..... 28
ตอนที่ 4 มโนทัศน์เกี่ยวกับทฤษฎีการสรุปอ้างอิง…….. 42
ตอนที่ 5 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ……. 49
ตอนที่ 6 งานวิจัยที่เกี่ยวข้อง..... 58
3 วิธีดำเนินการวิจัย..... 72
ตอนที่ 1 ประชากรและกลุ่มตัวอย่าง…….. 72
ตอนที่ 2 เครื่องมือที่ใช้ในการวิจัย..... 74
ตอนที่ 3 การสร้างเครื่องมือที่ใช้ในการวิจัย…… 74
ตอนที่ 4 การนำเครื่องมือที่ใช้ในการวิจัยไปเก็บรวบรวมข้อมูล……. 81
ตอนที่ 5 การวิเคราะห์ข้อมูล..... 82

4 ผลการวิเคราะห์ข้อมูล..... 87
ตอนที่ 1 ผลการศึกษาและพัฒนาวิธีการให้คะแนนการประเมินแผนผังมโนทัศน์. 87
ตอนที่ 2 ผลการเปรียบเทียบคุณภาพด้านความตรงและความเที่ยง..... 91
2.1 ผลการวิเคราะห์ข้อมูลเบื้องต้น……. 91
2.2 ผลการเปรียบเทียบความตรงตามสภาพระหว่างวิธีการตรวจให้คะแนนที่แตกต่างกัน..... 92
2.3 ผลการตรวจสอบความตรงเชิงโครงสร้างโดยใช้เมทริกซ์พหุลักษณะพหุวิธี……. 94
2.4 ผลการเปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบ ฟังก์ชันสารสนเทศของแบบสอบและประสิทธิภาพสัมพัทธ์เฉลี่ยระหว่างวิธีการตรวจให้คะแนนที่แตกต่างกัน..... 96
2.5 การวิเคราะห์คุณภาพด้านความเที่ยงโดยการวิเคราะห์ค่าสัมประสิทธิ์สรุปอ้างอิง ……..... 103
ตอนที่ 3 วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงเพื่อการตัดสินใจเชิงสรุปอ้างอิง 106
5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ..... 110
สรุปผลการวิจัย..... 110
อภิปรายผลการวิจัย..... 113
ข้อเสนอแนะ..... 118
รายการอ้างอิง..... 119
ภาคผนวก ก รายชื่อผู้เชี่ยวชาญ..... 126
ภาคผนวก ข ตารางแสดงแผนผังข้อสอบ..... 128
ภาคผนวก ค ค่าความยากและค่าอำนาจจำแนกของการวิเคราะห์ข้อสอบ..... 131
ภาคผนวก ง การวิเคราะห์ข้อมูลด้วยโปรแกรม TAP..... 134
ภาคผนวก จ การวิเคราะห์ข้อมูลด้วยโปรแกรม MULTILOG..... 136
ภาคผนวก ฉ การวิเคราะห์ข้อมูลด้วยโปรแกรม GENOVA..... 138
ภาคผนวก ช คู่มือการใช้แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์..... 141
ภาคผนวก ซ ตัวอย่างผลงานของนักเรียนและการตรวจให้คะแนน..... 154
ประวัติผู้เขียนวิทยานิพนธ์..... 157

สารบัญตาราง

ตาราง		หน้า
2.1	การแบ่งประเภทของแผนผังมโนทัศน์โดย Merle Tan และมนัส บุญประกอบ..	17
2.2	ส่วนประกอบของเมตริกซ์พหุลักษณะ – พหุวิธีขนาด 3 x 3.....	37
2.3	ประเภทของความตรง ความหมาย และวิธีการตรวจสอบ.....	40
2.4	สรุปผลการศึกษาในต่างประเทศและในประเทศเกี่ยวกับเทคนิคการใช้แผนผังมโนทัศน์และการให้คะแนนแผนผังมโนทัศน์.....	63
2.5	สรุปผลการนำแผนผังมโนทัศน์ไปใช้.....	65
3.1	ผลการตรวจสอบความตรงตามเนื้อหาโดยวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) ระหว่างเนื้อหาที่มุ่งวัดกับคำมโนทัศน์ที่ต้องการวัดในแผนผังเกณฑ์ของผู้ทรงคุณวุฒิ 3 ท่าน.....	78
4.1	ผลการตรวจสอบความตรงตามเนื้อหาโดยวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) ระหว่างตัวบ่งชี้กับคำคำอธิบายคุณภาพในเกณฑ์การตรวจให้คะแนนของผู้ทรงคุณวุฒิ 3 ท่าน…….	87
4.2	เกณฑ์การตรวจให้คะแนนแผนผังมโนทัศน์ของโนแวคและโกวิน.....	88
4.3	เกณฑ์การตรวจให้คะแนนแผนผังมโนทัศน์ของบอลท์.....	89
4.4	ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจำแนกตามวิธีการตรวจให้คะแนน.....	91
4.5	ค่า t ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนระหว่างวิธีการให้คะแนนที่ต่างกัน.....	92
4.6	ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากแบบทดสอบประเภทเลือกตอบ……	93
4.7	ความตรงตามสภาพจากการสอบจำแนกตามวิธีการตรวจให้คะแนน.....	93
4.8	ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากแบบทดสอบประเภทเลือกตอบ แผนผังมโนทัศน์ที่ตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินและตรวจให้คะแนนด้วยวิธีของบอลท์.....	94
4.9	เมทริกซ์พหุลักษณะพหุวิธี.....	95
4.10	ฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยที่ได้จากการตรวจให้คะแนนด้วยวิธีของโนแวค และโกวิน จำแนกตามผู้ประเมินและความสามารถของผู้สอบ.....	97
4.11	เปรียบเทียบความแตกต่างของค่าเฉลี่ยของฟังก์ชันสารสนเทศของข้อสอบเมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน โดยจำแนกตามผู้ประเมิน….	98

ตาราง หน้า

4.12 ฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยที่ได้จากการตรวจให้คะแนนด้วยวิธีของบอลท์ จำแนกตามผู้ประเมินและความสามารถของผู้สอบ..... 99
4.13 เปรียบเทียบความแตกต่างของค่าเฉลี่ยของฟังก์ชันสารสนเทศของข้อสอบเมื่อตรวจให้คะแนนด้วยวิธีของบอลท์ โดยจำแนกตามผู้ประเมิน..... 100
4.14 ฟังก์ชันสารสนเทศของข้อสอบเฉลี่ย จำแนกตามความสามารถของผู้สอบและวิธีการตรวจให้คะแนน..... 101
4.15 เปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบโดยเฉลี่ยจำแนกตามวิธีการตรวจให้คะแนน..... 102
4.16 ฟังก์ชันสารสนเทศของแบบสอบ จำแนกตามความวิธีการตรวจให้คะแนน..... 102
4.17 ค่าสารสนเทศเฉลี่ยของแบบสอบ จำแนกตามวิธีการตรวจให้คะแนน..... 103
4.18 อัตราส่วนสารสนเทศเฉลี่ยจำแนกตามวิธีการตรวจให้คะแนน..... 103
4.19 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมิน จำแนกตามผู้ประเมินและตามรูปแบบวิธีการให้คะแนน..... 104
4.20 ผลการเปรียบเทียบค่าความแปรปรวนและค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์เมื่อใช้รูปแบบวิธีการให้คะแนนของโนแวคและโกวิน (Novak & Gowin) ($n_p = 220,\ n_i = 2,\ n_r = 2$)..... 105
4.21 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง Coefficient เมื่อใช้วิธีการตรวจให้คะแนนของโนแวคและโกวิน และวิธีการตรวจให้คะแนนของบอลท์ กรณีมีผู้ประเมิน 1-6 คน ($n_p = 220,\ n_i = 4\text{-}28,\ n_r = 1\text{-}6$)..... 107

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

ภาพประกอบ หน้า

2.1 มโนทัศน์เกี่ยวกับมนุษย์ และองค์ประกอบ..... 15

2.2 แผนผังมโนทัศน์ย่อยเรื่องสัตว์..... 16

2.3 แนวคิดของการตรวจสอบความตรงตามสภาพ (ความตรงร่วมสมัย)..... 32

2.4 แนวคิดของการตรวจสอบความตรงเชิงทำนาย..... 33

2.5 มโนทัศน์ของการตรวจสอบความตรงเชิงทฤษฎี 34

2.6 แผนภูมิแสดงองค์ประกอบความแปรปรวนของรูปแบบ pxix r เมื่อตัวประกอบที่ศึกษาทั้งหมดเป็นตัวประกอบสุ่ม..... 46

2.7 โค้งฟังก์ชันสารสนเทศของ GRM..... 56

3.1 แสดงการตรวจผลงานนักเรียนของกลุ่มตัวอย่างผู้ประเมิน..... 73

3.2 การสร้างแบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์..... 75

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

เป้าหมายที่สำคัญประการหนึ่งของการจัดการศึกษาในทุกศาสตร์และทุกสาขาวิชาคือต้องการให้นักเรียนเกิดมโนทัศน์ในเนื้อหาที่เรียน เพราะถ้าหากผู้เรียนเกิดมโนทัศน์ก็สามารถทำให้เข้าใจจดจำเนื้อหาบทเรียนและสามารถสรุปสิ่งที่เรียนได้ดี นอกจากนี้ผู้เรียนยังสามารถสร้างองค์ความรู้และเชื่อมโยงความรู้ระหว่างสาขาวิชาได้ง่ายและรวดเร็วขึ้น ดังนั้นในปัจจุบันผู้สอนจึงมีวิธีการสอนที่หลากหลายโดยเฉพาะการสอนให้นักเรียนเกิดมโนทัศน์ ผู้เรียนจะต้องรู้จักเชื่อมโยงความรู้ และสร้างมโนทัศน์ด้วยตนเอง การเชื่อมโยงมโนทัศน์ของนักเรียนนั้นแสดงถึงความรู้ความเข้าใจในเนื้อหาวิชาและช่วยพัฒนาการคิดของนักเรียนเป็นอย่างดี แผนผังมโนทัศน์จึงถูกนำมาใช้กับผู้เรียน

จากการศึกษาเอกสารและงานวิจัยพบว่าในช่วงแรกๆ แผนผังมโนทัศน์ถูกนำมาใช้ในการจัดกิจกรรมการเรียนการสอน และต่อมาได้นำมาใช้เป็นแบบประเมินทางเลือกโดยสามารถประเมินนักเรียนได้ตามสภาพจริง จะเห็นว่าแผนผังมโนทัศน์ถูกนำมาใช้ใน 2 ลักษณะคือ ด้านการจัดการเรียนการสอนและด้านการประเมิน สำหรับในประเทศไทยส่วนใหญ่มีการศึกษาในด้านการจัดกิจกรรมการเรียนการสอน ส่วนในด้านการประเมินยังมีการศึกษาไว้น้อยมาก

การประเมินโดยใช้แผนผังมโนทัศน์ประกอบด้วย 3 ส่วนที่สำคัญคือ 1) งาน (task) ซึ่งเป็นหลักฐานที่แสดงถึงองค์ความรู้ของนักเรียนในเนื้อหาวิชานั้นๆ 2) รูปแบบการตอบของนักเรียน (response format) 3) ระบบการให้คะแนน (scoring system) ถ้าขาดองค์ประกอบใดองค์ประกอบหนึ่งจะทำให้แผนผังมโนทัศน์ขาดความน่าเชื่อถือในการประเมิน (Ruiz-Primo & Shavelson, 1996) และลักษณะของแผนผังมโนทัศน์ที่นำมาใช้มีหลากหลายรูปแบบซึ่งแต่ละรูปแบบมีลักษณะแตกต่างกันไป ความหลากหลายนำมาซึ่งความผันแปรสำหรับการใช้และการแปลความหมาย

จากการศึกษาในการประเมิน ปัญหาส่วนใหญ่ที่พบในการนำแผนผังมโนทัศน์ไปใช้คือการตรวจให้คะแนนแผนผังนั้นยังไม่มีความเป็นปรนัย เนื่องมาจากแผนผังมโนทัศน์ที่ซับซ้อนและยากต่อการให้คะแนนทำให้ผู้ตรวจให้คะแนนแต่ละคนให้คะแนนแตกต่างกัน และไม่สามารถตรวจให้ได้เหมือนเดิมในเวลาที่ต่างกัน และการให้คะแนนประพจน์ก็ไม่มีความเที่ยง ซึ่งในการประเมินแผนผังมโนทัศน์จะต้องพิจารณาในด้านความตรงและความเที่ยง ซึ่งเป็นคุณลักษณะสำคัญของการประเมิน

งาน (task) หรือสถานการณ์ที่กำหนดให้นักเรียนทำ จะมีแนวทางและวิธีการหาคำตอบหลายแนวทาง คำตอบที่ได้ก็อาจมิใช่เป็นไปตามแนวทางที่กำหนดไว้เสมอไป จึงทำให้การตรวจให้คะแนนไม่สามารถให้ได้อย่างชัดเจนแน่นอนเหมือนการตรวจให้คะแนนข้อสอบแบบเลือกตอบซึ่งอาจตรวจให้คะแนนด้วยเครื่องตรวจกระดาษคำตอบก็ได้ ดังนั้นการตรวจให้คะแนนแผนผังมโนทัศน์ที่ตัดสินใจโดยมนุษย์จึงต้องกำหนดแนวทางในการให้คะแนนไว้อย่างชัดเจน

สิ่งที่ใช้เป็นแนวทางในการประเมินแผนผังมโนทัศน์คือเกณฑ์การให้คะแนน ซึ่งมีลักษณะเป็นระดับของผลงาน โดยแต่ละระดับบรรยายถึงคุณภาพของผลงานในประเด็นที่ต้องการทำการประเมิน เป็นการตอบคำถามว่าผู้ถูกประเมินทำอะไรได้สำเร็จหรือว่ามีความสำเร็จขั้นใดจากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องพบว่า โนแวคและโกวิน (Novak & Gowin, 1984) และบอลท์ (Bolte, 1999) ได้เสนอวิธีการให้คะแนนแผนผังมโนทัศน์ ซึ่งมีรายละเอียดดังนี้

วิธีการให้คะแนนของโนแวคและโกวิน (Novak & Gowin, 1984) คือ แนวทางการให้คะแนนแผนผังมโนทัศน์ในวิชาคณิตศาสตร์โดยพิจารณาจากประพจน์ (propositions) การจัดลำดับขั้นตอน (hierarchy) การเชื่อมระหว่างมโนทัศน์ (cross link) และตัวอย่าง (examples)

วิธีการให้คะแนนของบอลท์ (Bolte, 1999) คือ แนวทางการให้คะแนนแผนผังมโนทัศน์แบบภาพรวม (holistic scoring) โดยแต่ละแผนผังมโนทัศน์มีคะแนนเต็ม 10 คะแนน โดยให้คะแนนของการจัดแผนผังมโนทัศน์ (organization) 6 คะแนน ความถูกต้อง (accuracy) 4 คะแนน

วิธีการให้คะแนนของโนแวคและโกวิน จะมีลักษณะการให้คะแนนแบบแยกองค์ประกอบเพื่อให้เห็นคุณภาพของการปฏิบัติงานรายละเอียดปลีกย่อย การประเมินจึงมีการวิเคราะห์แยกมิติหรือองค์ประกอบที่จะพิจารณาให้ระดับคะแนน ซึ่งแต่ละมิติหรือองค์ประกอบก็จะแบ่งระดับตามระดับคุณภาพความสามารถ วิธีการให้คะแนนของโนแวคและโกวินจะต้องรวมคะแนนเป็นคะแนนรวมทั้งหมดของผลงาน การตรวจให้คะแนนด้วยวิธีนี้จึงต้องใช้เวลามาก แต่รายละเอียดของแต่ละระดับจะมีประโยชน์ เมื่อต้องการที่จะวินิจฉัยข้อผิดพลาดของผู้เรียน (Kinchin et al, 2000; Klein, Chung, Osmundson, Herl, & O'Neil, 2002; McClure, Sonak, & Suen, 1999; XiuFeng & Hinchey, 1996) สำหรับวิธีการให้คะแนนของบอลท์ มีลักษณะการให้คะแนนแบบภาพรวม ซึ่งวิธีการตรวจให้คะแนนของบอลท์จะง่ายและประหยัดเวลา แต่จะไม่ได้รายละเอียดข้อผิดพลาดที่ชัดเจน (Ruiz-Primo, Schultz, Li, & Shavelson, 2001; Yin et al, 2004; รุ่งนภา ทศภานนท์, 2544; วิยะดา ระวังสุข, 2545)

การประเมินแผนผังมโนทัศน์เป็นกระบวนการที่มีความสำคัญมาก ซึ่งเป็นภาระที่ท้าทายสำหรับผู้ที่ใช้แผนผังในการประเมินผลการเรียนว่าจะมีวิธีการปฏิบัติอย่างไรในการพิจารณาว่าผลงานมีคุณภาพหรือไม่ โดยทั่วไปจะต้องพิจารณาในเรื่องความเที่ยง (reliability) และความตรง

(validity) กล่าวคือผลการประเมินมีความคงที่หรือไม่ และคะแนนที่ได้มีความหมายว่าอย่างไร คะแนนที่ได้สามารถสรุปพาดพิงไปสู่การตัดสินใจที่ถูกต้องเกี่ยวกับนักเรียน วิชา หรือโรงเรียนได้หรือไม่ ในการศึกษาว่าแผนผังมโนทัศน์นั้นมีคุณภาพด้านความเที่ยงหรือไม่ จะพิจารณาความสอดคล้องกันของผู้ประเมิน กล่าวคือ ผู้ประเมินทั้งหลายมีหน้าที่ตัดสินแผนผังมโนทัศน์ของนักเรียนต้องมีความสอดคล้องกันเกี่ยวกับคะแนนที่ให้แก่แผนผังมโนทัศน์ของนักเรียน และสอดคล้องกันในเรื่องวิธีประเมิน การให้คะแนนแผนผังมโนทัศน์ของนักเรียนของผู้ประเมินแต่ละคนในชิ้นงานเดียวกันควรใกล้เคียงกัน แต่ถ้าปราศจากซึ่งความสอดคล้องกันดังกล่าว แสดงว่าคะแนนที่ได้จากการประเมินจะขึ้นอยู่กับผู้ประเมินแต่ละคนไม่ได้ขึ้นอยู่กับคุณภาพของแผนผังมโนทัศน์ของนักเรียน (สมศักดิ์ ภู่วิภาดาวรรธน์, 2544)

การหาความเที่ยงของการปฏิบัติงาน โดยวิธีหาความสอดคล้องกันระหว่างผู้ประเมินเป็นที่ยอมรับในฐานะที่ใช้ตัดสินได้ว่าผลงานนั้นมีคุณภาพหรือไม่ เฮอร์แมนและวินเทอร์ (Herman and Winter, 1994) พบว่าแหล่งสำคัญของคะแนนความเที่ยงขึ้นอยู่กับสิ่งต่อไปนี้ 1) ความคงที่ของคะแนน แม้ว่าจะวัดเมื่อใดก็ตาม 2) ความคงที่ของคะแนนที่เกิดจากผู้ประเมินแต่ละคนในกลุ่มหรือคู่ของผู้ประเมิน และ 3) ผลงานหรือบริบท (context) ที่ได้รับการประเมิน

ส่วนมากวิธีการทั่วๆไปที่ใช้ในการประมาณค่าความเที่ยงระหว่างผู้ประเมินนั้น ได้มาจากการประยุกต์ใช้ทฤษฎีสรุปอ้างอิง (generalizability theory) อธิบายครั้งแรกโดยครอนบาคและคณะ (Cronbach et al, 1963) วิธีการนี้จะมีความยืดหยุ่นที่จะรวมแหล่งความคลาดเคลื่อนของการวัด รวมทั้งผู้ประเมิน ผลงาน จำนวนครั้ง และรูปแบบของการประเมิน ค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) สามารถแปลความเป็นดัชนีความเห็นพ้องต้องกันของผู้ประเมินโดยจะต้องเป็นผู้ประเมินเดียวกัน ฝึกในทิศทางเดียวกัน และตรวจให้คะแนนภายใต้เงื่อนไขเดียวกัน ความสอดคล้องระหว่างผู้ประเมินเป็นประเด็นที่สำคัญมากในการประเมินการปฏิบัติ ดังเช่น คอเรลซ์และคณะ (Korelz et al, 1994) แนะนำว่า ถึงแม้ว่าค่าความเที่ยงของผู้ประเมินจะมีขอบเขตในเรื่องของคะแนนจากทฤษฎีเทียบกับข้อเท็จจริงที่ได้จากการวัด ค่าความเที่ยงของผู้ประเมินที่สูงไม่ได้หมายความว่าคะแนนความเที่ยงจะเป็นที่น่าพอใจ ดังนั้นค่าความเที่ยงระหว่างผู้ประเมินจึงจำเป็น นอกจากนี้ ชัยพฤกษ์ เสรีรักษ์และคณะ (2540) กล่าวว่า การวิเคราะห์ความเที่ยงของการวัดผลงานของนักเรียนอย่างต่อเนื่องนั้นจะวิเคราะห์ด้วยทฤษฎีความเที่ยงแบบดั้งเดิม เช่น การแบ่งครึ่งข้อสอบแบบสัมประสิทธิ์แอลฟาของ ครอนบาค หรือใช้สูตรของ คูเดอร์ ริชาร์ดสัน ย่อมไม่เหมาะสมเพราะทฤษฎีเหล่านี้มีข้อตกลงว่า ข้อสอบวัดสิ่งเดียวกัน มีลักษณะคู่ขนานกัน และดำเนินการสอบครั้งเดียว แต่การประเมินด้วยแผนผังมโนทัศน์นี้เป็นการประเมินความคิดรวบยอดที่

มีลักษณะแตกต่างกัน จากทฤษฎีความเที่ยงที่เรียกว่า ทฤษฎีการสรุปอ้างอิง เป็นทฤษฎีที่เหมาะสมในการวิเคราะห์ค่าความเที่ยงของการประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์

ความตรงเชิงโครงสร้าง หรือความตรงเชิงทฤษฎี (construct validity) หมายถึง คุณสมบัติของแบบวัดที่ให้ผลการวัดสอดคล้องกับคุณลักษณะที่ต้องการวัด ซึ่งนิยามโดยตัวแปรโครงสร้างตามทฤษฎี ความตรงเชิงโครงสร้างเป็นความตรงที่มีความสำคัญมากที่สุด เพราะเป็นความตรงที่เชื่อมโยงการวัดในทางปฏิบัติกับลักษณะที่ต้องการวัดในทางทฤษฎี (นงลักษณ์ วิรัชชัย, 2537; วรรณี แกมเกตุ, 2539; Allen and Yen,1979; Kerlinger, 1986) หรือกล่าวอีกอย่างหนึ่งได้ว่าความตรงเชิงโครงสร้างเป็นคุณสมบัติที่สำคัญอย่างหนึ่งของแบบวัด ซึ่งเป็นตัวบ่งชี้ถึงคุณภาพของแบบวัดว่าสามารถวัดคุณลักษณะที่ต้องการวัดได้สอดคล้องตามโครงสร้างทฤษฎีหรือไม่

การตรวจสอบความตรงเชิงโครงสร้างเป็นกระบวนการ หรือกิจกรรมที่หาข้อมูลเชิงประจักษ์มาสนับสนุนสมมติฐานหรือโครงสร้างตามทฤษฎีที่ต้องการทดสอบ กล่าวคือ กระบวนการตรวจสอบความตรงเชิงโครงสร้าง จำเป็นต้องนิยามคุณลักษณะที่มุ่งวัดตามแนวคิดเชิงทฤษฎี ซึ่งเป็นนามธรรมให้อยู่ในรูปของตัวบ่งชี้หรือพฤติกรรมที่สามารถวัดได้ จากนั้นจึงนำผลการวัดเชิงประจักษ์มาตรวจสอบด้วยวิธีการต่างๆว่า สอดคล้องตามคุณลักษณะที่คาดหมายหรือสมมติฐานที่กำหนดไว้หรือไม่

สำหรับวิธีการในการการตรวจสอบเชิงโครงสร้างนั้นสามารถทำได้หลายวิธี เช่น ใช้หลักการวิเคราะห์เชิงตรรกะ วิธีการวิจัยเชิงทดลอง วิธีการศึกษาความสัมพันธ์ วิธีการเปรียบเทียบกับกลุ่มที่รู้จักหรือมีหลักฐานอยู่แล้ว วิธีการวิเคราะห์ด้วยทฤษฎีการสรุปอ้างอิง วิธีเมทริกซ์พหุลักษณะพหุวิธี วิธีการวิเคราะห์องค์ประกอบ รวมทั้งการวิเคราะห์โมเดลสมการโครงสร้าง (อุทุมพร จามรมาน, 2537) สำหรับในที่นี้จะกล่าวถึงเฉพาะวิธีเมทริกซ์พหุลักษณะพหุวิธี (multitrait-multimethod : MTMM) ซึ่งเป็นวิธีที่เหมาะสมกับสภาพการณ์การวิจัยทางสังคมศาสตร์และผู้วิจัยนำมาใช้ในการศึกษาครั้งนี้

จากการศึกษาเอกสารและงานวิจัยพบว่า รุซ-พรีโมและชาเวลสัน (Ruiz-Primo and Shavelson, 1996) ได้ยกประเด็นคำถามเกี่ยวกับความเที่ยงและความตรงของแผนผังมโนทัศน์ในฐานะของเครื่องมือสำหรับประเมิน ซึ่งต่อมาได้มีงานวิจัยที่ศึกษาติดตามมาเมื่อไม่นานมานี้ (Klein, Chung, Osmundson, Herl, & O'Neil, 2002; McClure, Sonak, & Suen, 1999; Nicoll, Francisco, & Nakhieh, 2001; Ruiz-Primo, Schultz, Li, & Shavelson, 2001; Ruiz-Primo, Shavelson, Li, &, Schultz 2001; Rye & Rubba, 2002) ซึ่งในการศึกษาวิจัยเหล่านี้ งานวิจัยบางเรื่องเปรียบเทียบงานของแผนผังมโนทัศน์ที่แตกต่างกัน (Ruiz-Primo, Shavelson, Li, &, Schultz 2001) งานวิจัยบางเรื่องวินิจฉัยระบบการให้คะแนนที่แตกต่างกัน (McClure, Sonak, & Suen, 1999) และงานวิจัยอื่นๆได้

ทดสอบความตรงของการประเมินโดยใช้แผนผังมโนทัศน์ โดยการใช้ Think-Aloud Protocols และการวัดอื่นๆ (Klein, Chung, Osmundson, Herl, & O'Neil, 2002; Ruiz-Primo, Schultz, Li, & Shavelson, 2001; Ruiz-Primo, Shavelson, Li, &, Schultz 2001)

จากความสำคัญของแผนผังมโนทัศน์และปัญหาเกี่ยวกับการตรวจให้คะแนนทำให้เห็นว่าการนำแผนผังมโนทัศน์มาใช้ในการประเมินด้วยวิธีการให้คะแนนที่มีคุณภาพจะสามารถทำให้ผู้เรียนฝึกคิดและพัฒนากระบวนการคิดได้อย่างเต็มศักยภาพ และทำให้ทราบถึงระดับความคิดรวบยอดและมโนทัศน์ที่คลาดเคลื่อนของนักเรียน ผู้วิจัยจึงสนใจที่ศึกษาคุณภาพของแผนผังมโนทัศน์ที่ใช้ในการประเมินเมื่อใช้วิธีตรวจให้คะแนนที่แตกต่างกัน โดยทำการเปรียบเทียบระหว่างวิธีการให้คะแนนของโนแวคและโกวิน (Novak & Gowin, 1984) และวิธีการให้คะแนนของบอลท์ (Bolte, 1999) ซึ่งผลที่ได้จากการศึกษาจะเป็นประโยชน์ต่อครูที่จะนำแผนผังมโนทัศน์มาประยุกต์ใช้เป็นแบบประเมินทางเลือกหนึ่ง ที่สามารถประเมินนักเรียนได้ตรงตามสภาพจริงและสามารถใช้และปรับปรุงพัฒนาผู้เรียนได้อย่างมีประสิทธิภาพต่อไป

คำถามวิจัย

จากการศึกษาประเด็นเกี่ยวกับวิธีการให้คะแนน ซึ่งมี 2 รูปแบบ ได้แก่ วิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ เมื่อนำมาใช้ประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์ทำให้เกิดคำถามวิจัยดังนี้

1. เกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ด้วยวิธีของโนแวคและโกวิน และวิธีของบอลท์เป็นอย่างไร

2. คุณภาพของแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์ เมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ จะมีค่าแตกต่างกันอย่างไร

3. จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสมเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ จะมีค่าแตกต่างกันอย่างไร

วัตถุประสงค์การวิจัย

1. เพื่อศึกษาและพัฒนาเกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ด้วยวิธีของโนแวคและโกวิน และวิธีของบอลท์

2. เพื่อเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ เมื่อตรวจโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และใช้วิธีการให้คะแนนของบอลท์ โดยพิจารณาจาก

2.1 ความตรงตามสภาพ

2.2 ความตรงเชิงโครงสร้าง

2.3 ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ

2.4 ความเที่ยงซึ่งแทนด้วยสัมประสิทธิ์สรุปอ้างอิง

3. เพื่อวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสมเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน หรือวิธีการให้คะแนนของบอลท์

ขอบเขตการวิจัย

1. เนื่องจากการประเมินครั้งนี้มีขั้นตอนการเก็บรวบรวมข้อมูลที่ต้องอาศัยความร่วมมือในการวิจัย คือ ให้ครูช่วยประเมินแผนผังมโนทัศน์ ดังนั้นผู้วิจัยจึงเลือกโรงเรียนที่ยินดีให้ความร่วมมือในจังหวัดนครราชสีมา จำนวน 1 โรงเรียน

2. ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นผู้ประเมินซึ่งเป็นครูสอนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1-3 และนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 ในโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน

กลุ่มตัวอย่าง ในการวิจัยครั้งนี้มีกลุ่มตัวอย่างที่ศึกษาประกอบด้วย

1. กลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 โรงเรียนสุรนารีวิทยา สำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต1 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ จำนวน 220 คน

2. กลุ่มตัวอย่างผู้ประเมิน เนื่องจากในการวิจัยครั้งนี้ต้องใช้ครูผู้สอนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1-3 เป็นผู้ประเมินแผนผังมโนทัศน์ซึ่งต้องใช้เวลาและความสนใจ ตั้งใจในการประเมิน ผู้วิจัยทำการเลือกครูผู้ประเมินโดยวิธีการแบบเจาะจง (purposive sampling) เพราะต้องหาผู้ประเมินที่เต็มใจและมีเวลาช่วยประเมินอย่างเต็มที่ มีประสบการณ์ในการสอนอย่างน้อย 3 ปี ครูที่ได้รับเลือกเป็นกลุ่มตัวอย่างรวมทั้งหมด 4 คน แบ่งครู 4 คน เป็น 2 กลุ่ม กลุ่มละ 2 คน โดยวิธีสุ่มอย่างง่ายโดยการจับสลาก ซึ่งผู้ประเมินกลุ่มที่ 1 จะใช้วิธีการให้คะแนนของโนแวคและโกวิน และผู้ประเมินกลุ่มที่ 2 จะใช้วิธีการให้คะแนนของบอลท์

3. ตัวแปรที่ใช้ในการวิจัย

3.1 ตัวแปรอิสระ ได้แก่ วิธีการให้คะแนนแผนผังมโนทัศน์

3.2 ตัวแปรตาม ได้แก่ ค่าความเที่ยงที่แทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง ฟังก์ชันสารสนเทศของข้อสอบ ความตรงตามสภาพและความตรงเชิงโครงสร้าง

คำจำกัดความที่ใช้ในการวิจัย

ความคิดรวบยอด หมายถึง ความรู้ ความเข้าใจที่เกิดจากกระบวนการคิดที่มีต่อเนื้อหาวิชาคณิตศาสตร์เรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ที่ได้รับจากการเรียนรู้หรือประสบการณ์โดยสรุปออกมาเป็นถ้อยคำ กลุ่มคำ หรือประโยคสั้นๆ กะทัดรัดที่สามารถสื่อความหมายได้

แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ หมายถึง เครื่องมือประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) เป็นแบบประเมินที่กำหนดคำมโนทัศน์มาให้ แล้วให้นักเรียนสร้างแผนผังลงในกระดาษที่แจกให้ โดยในการสร้างแผนผัง นักเรียนสามารถเพิ่มคำมโนทัศน์ได้ตามความคิดของตนเอง พร้อมทั้งลากเส้นเชื่อมโยงและเขียนคำเชื่อมแสดงความสัมพันธ์ระหว่างมโนทัศน์ให้สมบูรณ์มีจำนวน 2 แผนผัง แผนผังแบบมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) ดังกล่าวเป็นแบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์ เรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1

วิธีการให้คะแนนของโนแวคและโกวิน หมายถึง แนวทางการให้คะแนนแผนผังมโนทัศน์โดยพิจารณาจากประพจน์ (propositions) การจัดลำดับขั้นตอน (hierarchy) การเชื่อมระหว่างมโนทัศน์ (cross link) และตัวอย่าง (examples)

วิธีการให้คะแนนของบอลท์ หมายถึง แนวทางการให้คะแนนแผนผังมโนทัศน์แบบภาพรวม (holistic scoring) โดยแต่ละแผนผังมโนทัศน์มีคะแนนเต็ม 10 คะแนน โดยให้คะแนนของการจัดแผนผังมโนทัศน์ (organization) 6 คะแนน ความถูกต้อง (accuracy) 4 คะแนน

ผู้ประเมิน หมายถึง ครูผู้สอนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ถึง 3 มีประสบการณ์ด้านการสอนวิชาคณิตศาสตร์อย่างน้อย 3 ปี

ความเที่ยง หมายถึง คุณภาพของแผนผังมโนทัศน์ที่สามารถวัดมโนทัศน์ได้อย่างสม่ำเสมอ โดยพิจารณาจากค่าสัมประสิทธิ์การสรุปอ้างอิง และค่าฟังก์ชันสารสนเทศของข้อสอบ

สัมประสิทธิ์การสรุปอ้างอิง หมายถึง คุณภาพด้านความเที่ยงของคะแนนที่ได้จากการประเมินแผนผังมโนทัศน์โดยใช้ทฤษฎีสรุปอ้างอิง (Generalizability Theory) ซึ่งคำนวนจากการวิเคราะห์ความ แปรปรวนโดยใช้โปรแกรมสำเร็จรูป GENOVA

ฟังก์ชันสารสนเทศของข้อสอบ หมายถึง ดัชนีที่บ่งชี้คุณภาพของข้อสอบที่เกิดจากการผสมรวมของดัชนีคุณลักษณะของข้อสอบหลายลักษณะ ได้แก่ ค่าความยาก ค่าอำนาจจำแนกและค่าความแปรปรวนของคะแนนรายข้อ ซึ่งคำนวณได้จากการตรวจให้คะแนนแผนผังมโนทัศน์

ฟังก์ชันสารสนเทศของแบบสอบ หมายถึง ผลรวมของค่าสารสนเทศของข้อสอบทุกข้อในแบบสอบที่มีการตรวจให้คะแนนอย่างเดียวกันทั้งฉบับ โดยแสดงค่าฟังก์ชันสารสนเทศตามระดับความสามารถของผู้ตอบ ที่แตกต่างกัน ถ้าสูงที่ระดับใดแสดงว่ามีความแม่นยำสูงในการประมาณค่าพารามิเตอร์ของข้อกระทง ณ ระดับความสามารถนั้นๆ

ความตรง หมายถึง คุณภาพของแผนผังมโนทัศน์ที่สามารถวัดได้ตรงกับมโนทัศน์ที่มุ่งวัด โดยพิจารณาจากความตรง 2 ประเภท คือ

ความตรงตามสภาพ หมายถึง ความสอดคล้องระหว่างผลการประเมินโดยใช้แผนผังมโนทัศน์กับผลการทดสอบจากแบบทดสอบประเภทเลือกตอบวิชาคณิตศาสตร์ที่มุ่งวัดมโนทัศน์ในเรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1

ความตรงเชิงโครงสร้าง (construct validity) หมายถึงคุณสมบัติของแบบวัดที่ให้ผลการวัดสอดคล้องกับคุณลักษณะที่ต้องการวัด ได้จากการวิเคราะห์เมทริกซ์พหุลักษณะพหุวิธี (multitrait-multimethod : MTMM)

ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ในเชิงวิชาการ

1. ทราบวิธีการตรวจให้คะแนนแผนผังมโนทัศน์ที่ให้ค่าความตรงตามสภาพ และความตรงเชิงโครงสร้างสูง ซึ่งจะทำให้ผู้สอนเห็นคุณค่า และสนใจที่จะใช้แผนผังมโนทัศน์ในการวัดและประเมินผลมากขึ้น อันจะเป็นประโยชน์ต่อการพัฒนาศักยภาพของนักเรียนด้านความรู้ ความคิดตลอดจนสามารถสื่อสารความคิดในรูปการเขียนได้อย่างมีประสิทธิภาพ

2. ทราบค่าฟังก์ชันสารสนเทศของข้อสอบ เมื่อใช้วิธีการตรวจให้คะแนนแผนผังมโนทัศน์ที่แตกต่างกัน และสามารถนำไปใช้ประโยชน์ในการเลือกใช้แผนผังมโนทัศน์และวิธีการตรวจให้คะแนนที่เหมาะสมกับระดับความสามารถของกลุ่มตัวอย่าง เพื่อให้มีความแม่นยำในการประมาณค่าพารามิเตอร์ของผู้สอบ

3. ทำให้ทราบจำนวนผู้ประเมิน และจำนวนผลงานที่เหมาะสมสำหรับประเมินแผนผังมโนทัศน์เมื่อใช้วิธีการให้คะแนนที่แตกต่างกัน

ประโยชน์ในด้านการนำไปใช้

นักวัดผล และครูสามารถพิจารณาเลือกใช้วิธีการให้คะแนนแผนผังมโนทัศน์ จำนวนผู้ประเมิน และจำนวนแผนผังมโนทัศน์ที่เหมาะสมไปประยุกต์ใช้ในการประเมินผลการเรียนรู้ของผู้เรียนได้

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ผู้วิจัยแบ่งการนำเสนอออกเป็น 6 ตอนคือ

ตอนที่ 1 ความคิดรวบยอดทางคณิตศาสตร์

ความคิดรวบยอดและความคิดรวบยอดทางคณิตศาสตร์

ตอนที่ 2 แผนผังมโนทัศน์ (Concept Mapping)

ความหมายของแผนผังมโนทัศน์

ส่วนประกอบและลักษณะของแผนผังมโนทัศน์

เทคนิคแผนผังมโนทัศน์

การให้คะแนนแผนผังมโนทัศน์

ประโยชน์ของแผนผังมโนทัศน์

ตอนที่ 3 การตรวจสอบคุณภาพของเครื่องมือ

ความตรง

วิธีการตรวจสอบคุณภาพของความตรง

ตอนที่ 4 มโนทัศน์เกี่ยวกับทฤษฎีการสรุปอ้างอิง(Generalizability Theory)

แนวคิดพื้นฐานของทฤษฎีการสรุปอ้างอิง

การวิเคราะห์สัมประสิทธิ์สรุปอ้างอิง

ตอนที่ 5 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ (Item Response Theory)

ข้อตกลงเบื้องต้นเกี่ยวกับทฤษฎีตอบสนองข้อสอบ

มโนทัศน์เกี่ยวกับทฤษฎีตอบสนองข้อสอบแบบพหุวิภาค

ตอนที่ 6 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับแผนผังมโนทัศน์

งานวิจัยที่เกี่ยวข้องกับทฤษฎีการสรุปอ้างอิง

งานวิจัยที่เกี่ยวข้องกับทฤษฎีตอบสนองข้อสอบ

ตอนที่ 1 ความคิดรวบยอดทางคณิตศาสตร์

ความหมายของความคิดรวบยอด ได้มีนักการศึกษาให้ความหมายไว้ ดังนี้

คำว่า ความคิดรวบยอด หรืออาจใช้คำว่า มโนทัศน์ มโนมติ มโนภาพ ซึ่งในภาษาอังกฤษใช้คำว่า Concept สำหรับความหมายของมโนทัศน์นั้น มีนักศึกษาได้ให้ความหมายไว้ ดังนี้

กูดวิน และคลอสไมเออร์ (Goodwin and Klausmeier, 1975) ได้ให้ความหมายของความคิดรวบยอดไว้ว่า ความคิดรวบยอด หมายถึง ความสามารถในการเข้าใจเกี่ยวกับลักษณะของสิ่งหนึ่งๆ หรือหลายสิ่งไม่ว่าจะเป็นวัตถุ เหตุการณ์ หรือกระบวนการ ซึ่งทำให้สามารถแยกสิ่งต่างๆ ออกจากสิ่งอื่นได้ และในขณะเดียวกันก็สามารถจัดประเภทของสิ่งของประเภทเดียวกันไว้ด้วยกันได้

แมคคาวน์ และรุพ (McCown and Roup, 1992) ให้ความหมายของความคิดรวบยอด ไว้ว่า ความคิดรวบยอด หมายถึง ความคิดของบุคคลที่เกิดจากการเรียนรู้ การสังเกตหรือการสะสมประสบการณ์ โดยมโนทัศน์อาจเป็นวัตถุ เหตุการณ์ หรือความสัมพันธ์ที่มีลักษณะแตกต่างกันหรือเหมือนกันโดยสามารถสรุปรวมสิ่งต่างๆ เข้าด้วยกัน และสามารถแยกแยะความแตกต่างออกจากกันได้

เพรสลีและแมคคอร์มิก (Pressley and McComick, 1995) ให้ความหมายของความคิดรวบยอด ไว้ว่า ความคิดรวบยอด หมายถึง ความเข้าใจของบุคคลเกี่ยวกับกลุ่มของสิ่งเร้าทีมีความสัมพันธ์กัน โดยที่กลุ่มของสิ่งเร้าอาจเป็นวัตถุ เหตุการณ์ หรือกระบวนการก็ได้ ซึ่งมีลักษณะบางประการร่วมกัน

สุรพันธ์ ต้นศรีวงษ์ (2538) ได้ให้ความหมายของความคิดรวบยอดไว้ว่า ความคิดรวบยอด คือ แนวความคิดของบุคคลที่มีต่อสิ่งต่างๆ ซึ่งทำให้บุคคลนั้นสามารถสรุปรวมลักษณะเหมือนหรือแยกแยะลักษณะแตกต่างของคุณสมบัติของสิ่งของหรือเหตุการณ์นั้น ๆ

สุชา จันทร์เอม (2540) ได้ให้ความหมายของความคิดรวบยอดไว้ว่า ความคิดรวบยอด หมายถึง ความเข้าใจของคนที่มีต่อสิ่งต่างๆ ซึ่งแสดงออกในรูปของสัญลักษณ์ที่ใช้แทนสิ่งของหรือสถานการณ์หลาย ๆ อย่างที่มีความหมายร่วมกัน อยู่อย่างหนึ่ง เป็นต้นว่า เมื่อนึกถึงปลาวาฬ สุนัข วัว ควาย จะเห็นว่า ลักษณะร่วมกันอย่างหนึ่งของสัตว์เหล่านี้ คือ มีน้ำนมเลี้ยงลูกอ่อน ซึ่งเรียกสัตว์เหล่านี้ว่า สัตว์เลี้ยงลูกด้วยนม ซึ่งเป็นความคิดรวบยอดที่รวมเอาสัตว์ที่มีรูปร่างแตกต่างกันเข้าไว้ด้วยกัน และแยกสัตว์พวกนี้ออกจากสัตว์กลุ่มอื่น

อาภรณ์ ใจเที่ยง (2540) ให้ความหมายของความคิดรวบยอดไว้ว่า ความคิดรวบยอด คือ การจัดลักษณะที่เหมือน ๆ กัน ของประสบการณ์ หรือสิ่งของเข้าด้วยกันอย่างมีระเบียบ ทำให้เกิดเป็นหน่วยของความคิดหรือประเภทของประสบการณ์ อาจกล่าวได้ว่า ภายในช่วงระยะเวลาหนึ่ง

และความคิดรวบยอดนี้อาจเปลี่ยนแปลงได้ เมื่อผู้เรียนมีประสบการณ์มากขึ้นหรือมีวุฒิภาวะเพิ่มขึ้น

จากความหมายของความคิดรวบยอดดังกล่าวข้างต้น สามารถสรุปความหมายของความคิดรวบยอดได้ว่า ความคิดรวบยอดหมายถึง ความคิด ความเข้าใจของบุคคลที่มีต่อสิ่งต่างๆ โดยอาจเป็นวัตถุเหตุการณ์หรือกระบวนการ โดยที่บุคคลสามารถสรุปลักษณะร่วมที่เหมือนกันหรือแยกแยะลักษณะที่แตกต่างกันได้

ความหมายของความคิดรวบยอดทางคณิตศาสตร์ ได้มีนักการศึกษาให้ความหมายไว้ ดังนี้

โดโนแวน และเจอรอลด์ (Donovan and Gerald, 1972) ให้ความหมายของความคิดรวบยอดทางคณิตศาสตร์ไว้ว่า ความคิดรวบยอดทางคณิตศาสตร์ หมายถึง ความคิดของบุคคลซึ่งเป็นนามธรรมเกี่ยวกับความรู้ทางคณิตศาสตร์ เช่น สมบัติของวัตถุ หรือ เหตุการณ์ต่างๆ โดยสามารถบอกลักษณะร่วมและลักษณะที่แตกต่างของแต่ละความคิดรวบยอดได้ เช่น ความคิดรวบยอด “สาม” เป็นความคิดรวบยอดที่เป็นนามธรรมที่ใช้แทนความหมายของสิ่งของสามสิ่ง

คูนี และ เฮนเดอร์สัน(Cooney and Henderson, 1975) ได้ให้ความหมายของความคิดรวบยอดทางคณิตศาสตร์ ไว้ว่า ความคิดรวบยอดทางคณิตศาสตร์ หมายถึง ความเข้าใจเกี่ยวกับวิชาคณิตศาสตร์ที่ได้เรียนรู้ โดยนักเรียนสามารถสรุปความเข้าใจที่ได้ออกมาในรูปของนิยามหรือความหมายของเรื่องนั้น เช่น การมีความคิดรวบยอดทางคณิตศาสตร์เรื่องฟังก์ชัน คือนักเรียนสามารถบอกนิยามของฟังก์ชันได้

เอ็กเจนและโคแซก (Eggen and Kauchak, 1995) ได้ให้ความหมายของความคิดรวบยอดทางคณิตศาสตร์ไว้ว่า ความคิดรวบยอดทางคณิตศาสตร์ หมายถึง ความคิด ความเข้าใจของบุคคลที่มีต่อสิ่งเร้า ซึ่งบุคคลสามารถจัดประเภทหรือจัดกลุ่มของสิ่งเร้าที่มีคุณสมบัติบางประการร่วมกันโดยผ่านกระบวนการเรียนรู้ เช่น ความคิดรวบยอดของรูปสี่เหลี่ยม ผืนผ้า คือ รูปสี่เหลี่ยมที่มีขนาดของมุมทั้งสี่เท่ากัน และเท่ากับ 90 องศา มีด้านตรงข้ามเท่ากันและขนานกัน เป็นต้น

พรรณทิพย์ ม้ามณี (2520) ได้ให้ความหมายของความคิดรวบยอดทางคณิตศาสตร์ไว้ว่า ความคิดรวบยอดทางคณิตศาสตร์ หมายถึง ความเข้าใจและความสามารถในการเก็บใจความ หรือย่อเนื้อหาที่เรียนได้ รวมทั้งสามารถนำเอาไปใช้หรือสร้างเป็นกรณีทั่วไปได้ ซึ่งเป็นความหมายที่กว้างกว่าความเข้าใจธรรมดา

จากความหมายดังกล่าวข้างต้น สามารถสรุปความหมายของความคิดรวบยอดทางคณิตศาสตร์ได้ว่า ความคิดรวบยอดทางคณิตศาสตร์ หมายถึง ความคิด ความเข้าใจของบุคคลเกี่ยวกับคณิตศาสตร์อันเกิดจากการได้เรียนรู้หรือรับประสบการณ์ในการเรียนคณิตศาสตร์ โดยการสรุปความเข้าใจที่ได้ออกมาเป็นนิยาม สามารถจัดประเภทของสิ่งเร้าที่เหมือนกันเข้าด้วยกัน และแยกแยะความแตกต่างของสิ่งเร้าที่ไม่สัมพันธ์กันออกจากกัน

ตอนที่ 2 แผนผังมโนทัศน์ (Concept Mapping)

ความหมายของแผนผังมโนทัศน์

แผนผังมโนทัศน์ไม่ใช่เรื่องใหม่ของวงการศึกษาเพราะมีผู้ริเริ่มใช้มานานกว่า 25 ปี (Zeilik, 2000) คำว่าแผนผังมโนทัศน์ตรงกับคำอื่นๆในภาษาอังกฤษอีกหลายคำ เช่น Concept maps หรือ C-Maps, Conceptual Framework, Semantic Mapping, Semantic Maps, Plot Maps, Concept webs เป็นต้น ส่วนคำในภาษาไทยพบว่ามีความแตกต่างกันหลายคำ เช่น กรอบมโนคติ กรอบมโนทัศน์ ผังมโนทัศน์ ผังมโนภาพ แผนที่มโนมติ แผนผังมโนมติ เป็นต้น (มนัส บุญประกอบ, 2543)

แผนผังมโนทัศน์เกิดขึ้นจากการพัฒนาของ โจเซฟ ดี โนแวค ภายใต้ทฤษฎีของออซูเบล ซึ่งเป็นทฤษฎีที่ว่าด้วยผู้เรียนเป็นศูนย์กลางในการสร้างความหมายหรือการสร้างความรู้ การเรียนรู้อย่างมีความหมายจะปรากฏขึ้นเมื่อผู้เรียนดูดซึมข้อมูลใหม่เข้าสู่โครงสร้างทางปัญญาที่มีอยู่ โดยโครงสร้างทางปัญญาของบุคคลนั้นมีลักษณะเป็นลำดับขั้นลดหลั่นกันลงมา (hierarchy) โดยที่โนแวค ทำให้แนวคิดของออซูเบลได้มองเห็นเป็นแผนภูมิที่เป็นรูปธรรมได้ และได้วิจัยบุกเบิกอย่างจริงจังต่อเนื่องมาตั้งแต่ ค.ศ. 1973 จากการศึกษาพบว่ามีนักการศึกษาหลายท่านที่สนใจเกี่ยวกับแผนผังมโนทัศน์และให้ความหมายต่างๆกันออกไป (Gowin and Johensen, 1983; Novak, 1984; Novak, 1993; Zeilik,2000; and Hin-wai, n.d.) ได้ให้ความหมายคล้ายคลึงกันว่า แผนผังมโนทัศน์เป็นแผนภาพที่แสดงความสัมพันธ์ระหว่างมโนทัศน์หลายมโนทัศน์โดยมีคำหรือข้อความเพื่อแสดงถึงความสัมพันธ์ ซึ่งจะเริ่มเขียนมโนทัศน์หลักให้อยู่ด้านบนสุด จากนั้นเขียนมโนทัศน์รองลดหลั่นกันลงมา และมีการลากเส้นเชื่อมโยงความสัมพันธ์ระหว่างมโนทัศน์และเขียนคำหรือข้อความเพื่อแสดงความสัมพันธ์ระหว่างมโนทัศน์ไว้ที่เส้น นอกจากนั้นมีนักการศึกษาอีกหลายท่านที่ให้นิยามและความหมายของแผนผังมโนทัศน์ไว้ดังนี้

ฮาร์ท (Hart, 1994) และ ฟราเซอร์ (Fraser, 1996) ได้ให้ความหมายแผนผังมโนทัศน์ไว้คล้ายคลึงกันว่า เป็นเครื่องมือที่ใช้ตรวจสอบความเข้าใจและส่งเสริมกระบวนการคิดและการเรียนรู้ของนักเรียน ซึ่งลักษณะของแผนผังมโนทัศน์ เป็นกลุ่มคำหรือข้อมูลสารสนเทศโดยนักเรียนเป็นผู้สร้างขึ้น เพื่อนำเสนอความรู้ความเข้าใจเกี่ยวกับความคิดรวบยอดและความสัมพันธ์ของความคิด

ทรอคิม (Trochim, 1997) ได้ให้ความหมายว่า แผนผังมโนทัศน์เป็นโครงสร้างของแนวคิด (conceptualization) โดยกลุ่มคนที่ต้องการพัฒนากรอบมโนทัศน์สำหรับเป็นแนวทางในการวางแผนหรือการประเมินผล ซึ่งแผนผังมโนทัศน์จะเอื้อให้กลุ่มตีความในกิจกรรมต่างๆได้รวดเร็ว เนื่องจากแสดงออกมาเป็นกราฟฟิคหรือรูปภาพโดยแสดงกลุ่มมโนทัศน์ทั้งหมดและความสัมพันธ์ระหว่างมโนทัศน์เหล่านั้นได้พร้อมๆกันหรือในเวลาเดียวกัน

แลนซิง (Lanzing, 2002) ได้ให้ความหมายของแผนผังมโนทัศน์ว่าเป็นเทคนิคการเป็นตัวแทนองค์ความรู้ในรูปกราฟ ซึ่งเป็นเครือข่าย (networks) ของมโนทัศน์ เครือข่ายนั้นแสดงออกมาในรูปของกลุ่มคำ (node) และการเชื่อมโยง (link) ซึ่งอาจเป็นเส้นตรงหรือเส้นโค้งโดยกลุ่มคำจะเป็นตัวแทนของมโนทัศน์ และการเชื่อมโยงจะเป็นตัวแทนของความสัมพันธ์ระหว่างมโนทัศน์เหล่านั้น

รุซ-พรีโม และคณะ (Ruiz-Primo et al, 2001) กล่าวว่า แผนผังมโนทัศน์ หมายถึงแผนภาพที่มีส่วนแสดงมโนทัศน์ เส้นที่เชื่อมโยงระหว่างมโนทัศน์เพื่อแสดงว่ามโนทัศน์มีความสัมพันธ์กัน และคำเชื่อมโยงบนเส้นเป็นการแสดงว่ามโนทัศน์นั้นมีความสัมพันธ์กันอย่างไร

นักการศึกษาไทยหลายท่าน (กำพล ดำรงวงศ์, 2540; มนัส บุญประกอบ, 2542; และสาลิกา เมธนาวิน, 2545; วิยะดา ระวังสุข, 2545) ได้ให้ความหมายของแผนผังมโนทัศน์ไว้ว่า เป็นแผนภูมิที่แสดงถึงความสัมพันธ์อย่างมีความหมายของมโนทัศน์อย่างมีลำดับขั้นตอนและระหว่างมโนทัศน์มีการเชื่อมโยงด้วยเส้นตรงหรือเส้นโค้งและบางครั้งมีการระบุคำหรือประพจน์กำกับไว้บนเส้นโยงระหว่างคำมโนทัศน์นั้นด้วย

จากการศึกษาสามารถสรุปความหมายของแผนผังมโนทัศน์ได้ว่า เป็นเครื่องมือประเมินประเภทหนึ่งที่สามารถตรวจสอบกระบวนการคิดและเสนอความคิดอย่างมีลำดับขั้นตอน โดยมีลักษณะเป็นแผนภาพที่แสดงถึงมโนทัศน์ต่างๆหลายมโนทัศน์ ซึ่งเริ่มเขียนมโนทัศน์หลักไว้ด้านบนแล้วเขียนมโนทัศน์รองและ มโนทัศน์ย่อยลดหลั่นกันลงมา ระหว่างมโนทัศน์มีการเชื่อมโยงด้วยเส้นตรงหรือเส้นโค้ง (linking line) และเขียนคำหรือวลี (linking Phrase)ไว้บนเส้นเพื่อแสดงถึงความสัมพันธ์ระหว่างมโนทัศน์

ส่วนประกอบและลักษณะของแผนผังมโนทัศน์

ส่วนประกอบของแผนผังมโนทัศน์นั้น นักการศึกษากล่าวถึงดังนี้ โนแวค (Novak, 1991) กล่าวไว้ว่าแผนผังมโนทัศน์ประกอบด้วยส่วนประกอบสำคัญ 4 ส่วน ได้แก่

1. มโนทัศน์ (concept) หมายถึง คำที่ใช้แทนชื่อของมโนทัศน์ เป็นคำ หรือวลีสั้นๆ

2. ความสัมพันธ์ (relationship) หรือการเชื่อมโยงระหว่างประพจน์ (propositional linkage) เป็นการลากเส้นเชื่อมโยงระหว่างคำมโนทัศน์

3. ลำดับขั้น (hierarchy) เป็นชั้นหรือลำดับขั้นของความสัมพันธ์ระหว่างมโนทัศน์หลักไปยังมโนทัศน์รอง หรือมโนทัศน์ย่อย

4. การเชื่อมโยงข้ามมโนทัศน์ (cross-links) ซึ่งเป็นการแสดงถึงความสัมพันธ์เชื่อมโยงข้ามชั้นหรือข้ามขั้นระหว่างมโนทัศน์ จากมโนทัศน์หนึ่งไปอีกมโนทัศน์หนึ่ง

และมีนักการศึกษาต่างประเทศอีกหลายท่าน (Plotnick, 1997; Baroody and Bartels, 2000; Classidy, Griffiths and Nakoechny, 2001) ได้กล่าวถึงองค์ประกอบของแผนผังมโนทัศน์ไว้ว่า แผนผังมโนทัศน์จะต้องมีส่วนประกอบ 3 ส่วนที่สำคัญดังนี้

1. คำมโนทัศน์ (concept) เป็นชื่อมโนทัศน์ของเรื่องที่นำมาสร้างแผนผังมโนทัศน์ซึ่งอาจจะเป็นคำหรือวลีหรือประโยคสั้นๆโดยจะเขียนไว้ในกรอบ ซึ่งอาจเป็นรูปวงกลมหรือวงรี หรือสี่เหลี่ยมก็ได้

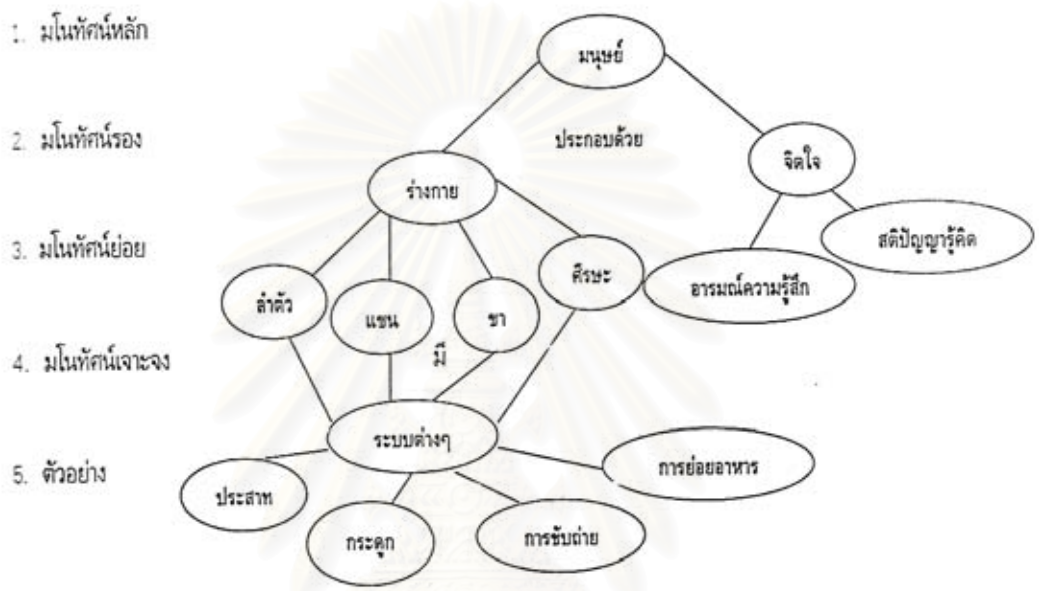
2. เส้นเชื่อมโยงระหว่างมโนทัศน์ (linking lines) เป็นเส้นที่ลากเชื่อมโยงระหว่างคำมโนทัศน์สองมโนทัศน์ โดยเส้นเชื่อมโยงความสัมพันธ์นี้อาจจะเป็นเส้นตรง เส้นโค้ง บางครั้งอาจแสดงทิศทางความสัมพันธ์ด้วยหัวลูกศรทางเดียว หรือสองทิศทางกำกับไว้ด้วย

3. คำเชื่อม (linking phrases) เป็นคำที่แสดงถึงความสัมพันธ์ระหว่างมโนทัศน์สองมโนทัศน์ โดยคำเชื่อมโยงนั้นจะต้องเป็นคำที่ทำให้เกิดประพจน์ที่มีความหมาย อาจเป็นถ้อยคำ วลี หรือประโยค ซึ่งต้องเป็นคำที่สั้น กะทัดรัด ทำให้ผู้อ่านเข้าใจง่าย

จากส่วนประกอบของแผนผังมโนทัศน์ที่กล่าวมาข้างต้น สามารถสรุปส่วนประกอบที่สำคัญของแผนผังมโนทัศน์ได้ดังนี้คือ คำมโนทัศน์ (concept) คำเชื่อม (linking phrases) และเส้นเชื่อมระหว่างมโนทัศน์ (linking line) นอกจากนั้นสามารถเพิ่มเติมการเชื่อมโยงข้ามสาย (cross links) ได้ เพื่อเป็นการแสดงถึงความเข้าใจในความสัมพันธ์ระหว่างมโนทัศน์

ลักษณะของแผนผังมโนทัศน์

ลักษณะของแผนผังมโนทัศน์มีองค์ประกอบหลายองค์ประกอบ โนแวค (Novak, 1990) ได้เสนอลักษณะแผนผังอย่างง่ายที่มุ่งแสดงความสัมพันธ์ระหว่างมโนทัศน์ระดับต่างๆไว้ดังแผนภาพที่ 2.1



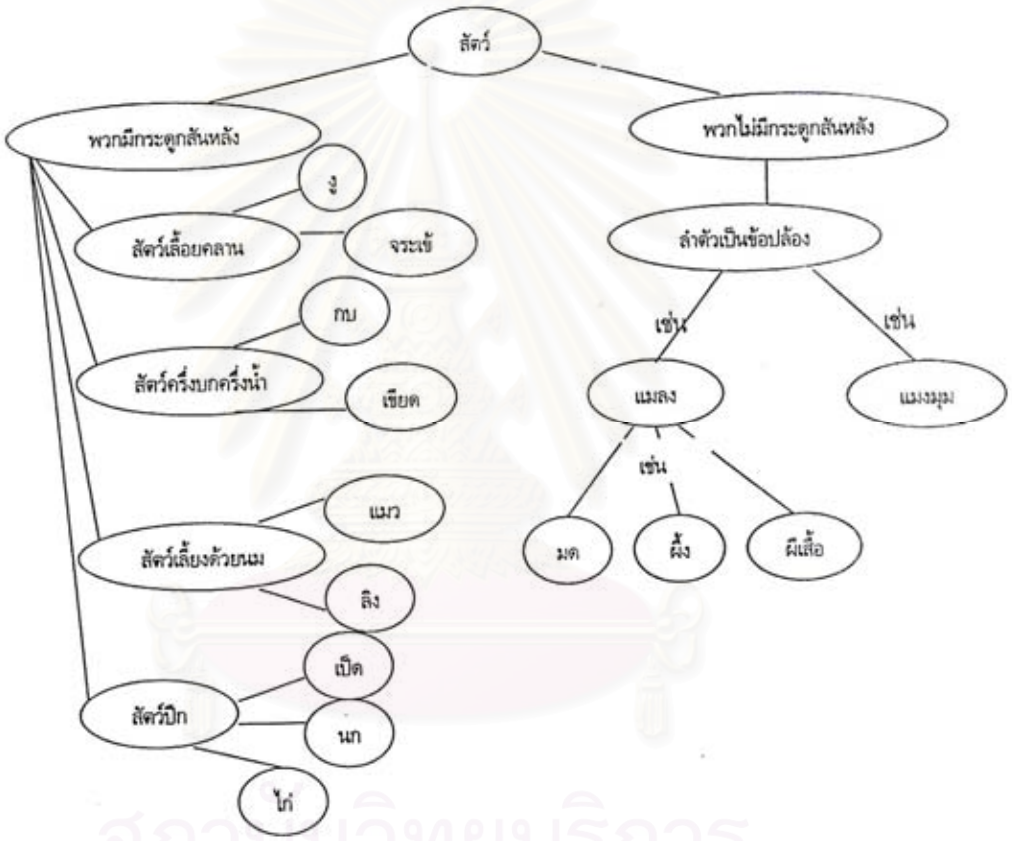
แผนภาพที่ 2.1 มโนทัศน์เกี่ยวกับมนุษย์ และองค์ประกอบ (Novak, 1984 อ้างถึงใน มนัส บุญประกอบ, 2542)

ประเภทของแผนผังมโนทัศน์ มีหลายรูปแบบซึ่งมีชื่อเรียกแตกต่างกันไป สำหรับการนำแผนผังมโนทัศน์มาใช้ก็ขึ้นอยู่กับลักษณะข้อมูลหรือวัตถุประสงค์ของการศึกษา นักการศึกษาหลายท่านได้จัดแบ่งประเภทของแผนผังมโนทัศน์โดยใช้เกณฑ์ต่างกันไปหลายแนวคิด โดยมีลักษณะที่หลากหลายดังนี้

มินท์เทล วอนเดอร์ซี และโนแวค (Mintzes, Wandersee, and Novak, 1998 อ้างถึงใน มนัส บุญประกอบ, 2542) ได้แบ่งแผนผังมโนทัศน์ออกเป็น 2 ชนิด ได้แก่

1. แผนผังหลัก (Macro map) ซึ่งแสดงความสัมพันธ์ระหว่างมโนทัศน์ที่สำคัญๆเท่านั้น

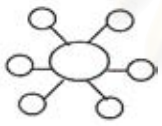
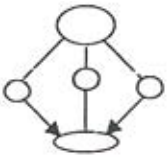
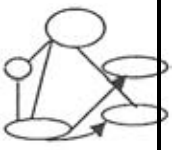
2. แผนผังย่อย (Micro map) แสดงรายละเอียดเพิ่มเติม เฉพาะมโนทัศน์องค์ประกอบของแผนผังหลัก ดังรูป



แผนภาพที่ 2.2 แผนผังมโนทัศน์ย่อยเรื่องสัตว์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เมิร์ล ตัน (Merle Tan, 2532 อ้างถึงใน มนัส บุญประกอบ, 2543) แห่งมหาวิทยาลัยฟิลิปปินส์ ได้แบ่งแผนผังมโนทัศน์ออกเป็น 4 ประเภท ได้แก่

ประเภท	1. กระจายออก	2.ปลายเปิด	3. เชื่อมโยงข้ามชุด	4. ปลายปิด	5. ผสม
ลักษณะ	มีการกระจายเชื่อมโยงคำมโนทัศน์รอบๆบางแห่งเรียกว่า Mind Mapping หรือ Spider Map	การเชื่อมโยงระหว่างมโนทัศน์จะลดหลั่นกันไปทางด้านล่าง	คล้ายกับแบบปลายเปิดแต่สามารถเชื่อมโยงข้ามกลุ่มทิศทางได้	คล้ายแบบปลายเปิดแต่มีการกำหนดมโนทัศน์ด้านล่างสุดปิดการเชื่อมโยง	มีโครงสร้างต่างชนิดปนอยู่ตั้งแต่ 2 ชนิดขึ้นไป
โครงสร้าง					

ตารางที่ 2.1 การแบ่งประเภทของแผนผังมโนทัศน์โดย Merle Tan และมนัส บุญประกอบ

นอกจากนี้ คลาสสิดี กริฟฟริท และแนคคอชนี (Classidy, Griffiths and Nakoechny ,2001) ได้จำแนกแผนผังมโนทัศน์ออกเป็น 3 ประเภทดังนี้

1. แผนผังต้นไม้ โดยมโนทัศน์จะถูกจัดเป็นหมวดหมู่ การเขียนมโนทัศน์จะเริ่มจากส่วนบนของกระดาษ และเขียนมโนทัศน์อื่นๆลดหลั่นกันลงไป

2. แผนผังที่มีลักษณะเป็นเว็บ เป็นแผนผังที่ไม่มีลำดับขั้นตอน ซึ่งสามารถเขียนประพจน์เชื่อมกันได้

3. แผนผังแบบวง แผนผังมโนทัศน์ประเภทนี้จะเขียนมโนทัศน์ไว้ตรงกลางแล้วเขียนประพจน์ออกจากจุดศูนย์กลางออกไปด้านข้าง

จากที่กล่าวมาจะเห็นว่าแผนผังมโนทัศน์มีหลายประเภทจะมีการนำเสนอที่แตกต่างกันไป การใช้ก็ขึ้นอยู่กับผู้เลือกว่าจะใช้ประเภทใดจึงจะเหมาะสมกับเรื่องที่ตนเองสนใจ ซึ่งสามารถสรุป

ประเภทของแผนผังมโนทัศน์ได้ 3 ประเภท คือ 1) แผนผังที่มีลำดับขั้นตอน 2) แผนผังที่ไม่มีลำดับขั้นตอน และ 3) แผนผังที่มีลักษณะเป็นแบบวง

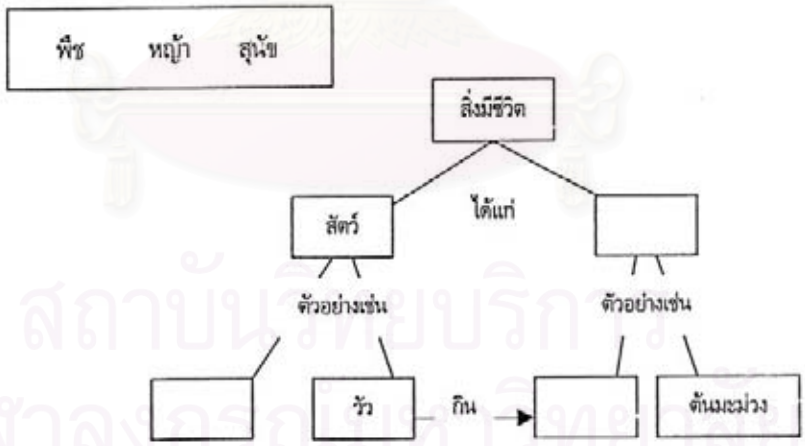
เทคนิคแผนผังมโนทัศน์

เทคนิคแผนผังมโนทัศน์มีหลายประเภท ในการนำมาใช้ก็ขึ้นอยู่กับวัตถุประสงค์ของการวิจัย ซึ่งนักการศึกษาหลายท่านได้นำเสนอไว้ดังนี้

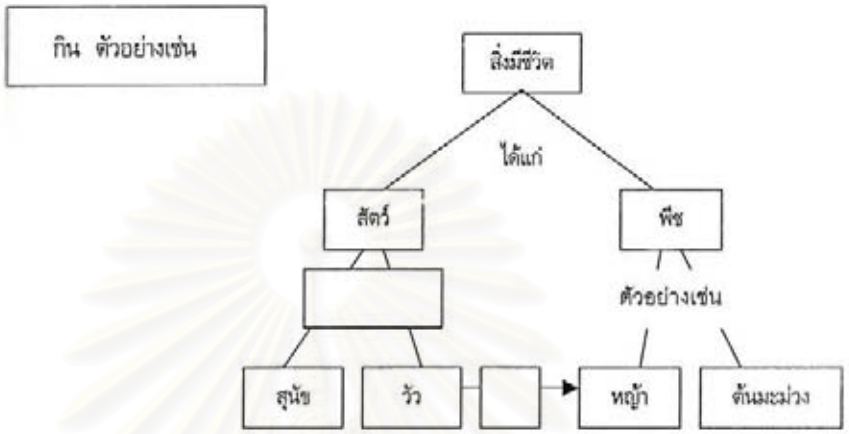
รุซ-พรีโม และคณะ (Ruiz-Primo et al, 2001) ได้ศึกษาเปรียบเทียบประเภทของเทคนิคแผนผังมโนทัศน์ไว้ 2 ประเภท คือ

1. เทคนิคแผนผังมโนทัศน์แบบเติมคำลงในผัง (fill-in-the-map) เป็นเทคนิคการจัดแผนผังมโนทัศน์ที่มีการกำหนดคำมโนทัศน์ทั้งหมดมาให้ โดยกำหนดเป็นแผนผังโครงสร้างที่ยังไม่สมบูรณ์แล้วให้ นักเรียนนำมโนทัศน์ที่กำหนดให้ทั้งหมดมาเติมลงในช่องว่างให้สมบูรณ์ ซึ่งแบ่งเป็น 2 ลักษณะคือ

1.1 แบบเติมคำมโนทัศน์ (fill-in-the-nodes) เป็นการจัดแผนผังมโนทัศน์โดยการเติม มโนทัศน์ในช่องว่างของแผนผังโครงสร้างให้สมบูรณ์ ดังตัวอย่าง



1.2 แบบเติมคำเชื่อมบนเส้น (fill-in-the-lines) เป็นการจัดแผนผังมโนทัศน์ โดยการเติมคำเชื่อมลงบนเส้นโยงความสัมพันธ์ระหว่างมโนทัศน์ที่ว่างให้สมบูรณ์ ดังตัวอย่าง

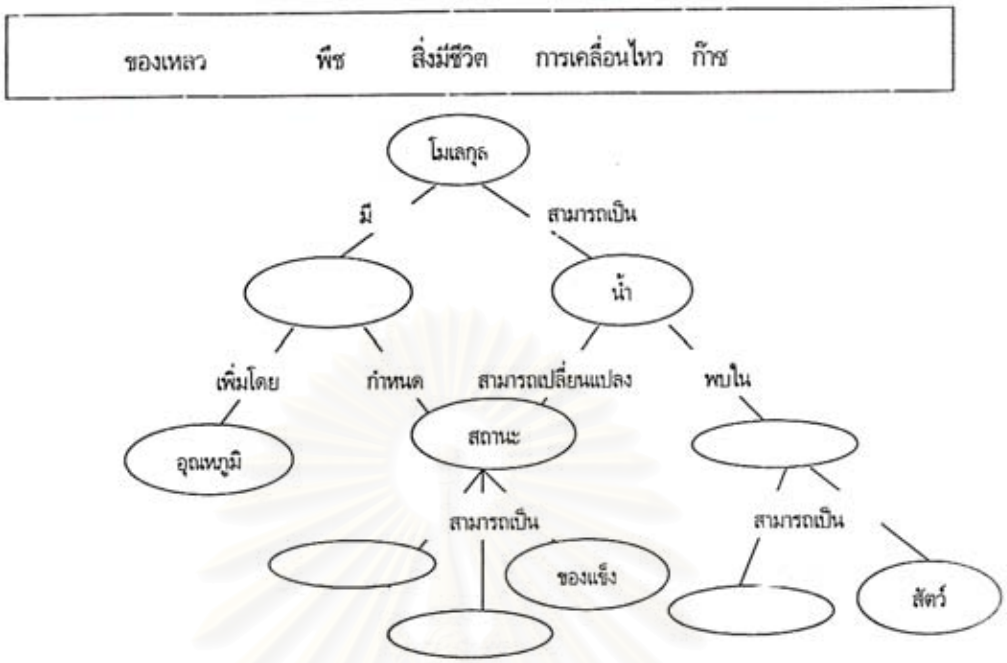


2. เทคนิคการจัดแผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) เป็นการจัดแผนผังมโนทัศน์โดยกำหนดรายการของมโนทัศน์บางส่วนมาให้ โดยให้นักเรียนนำคำมโนทัศน์เหล่านั้นมาสร้างเป็นแผนผังมโนทัศน์ให้สมบูรณ์ ตัวอย่างมโนทัศน์

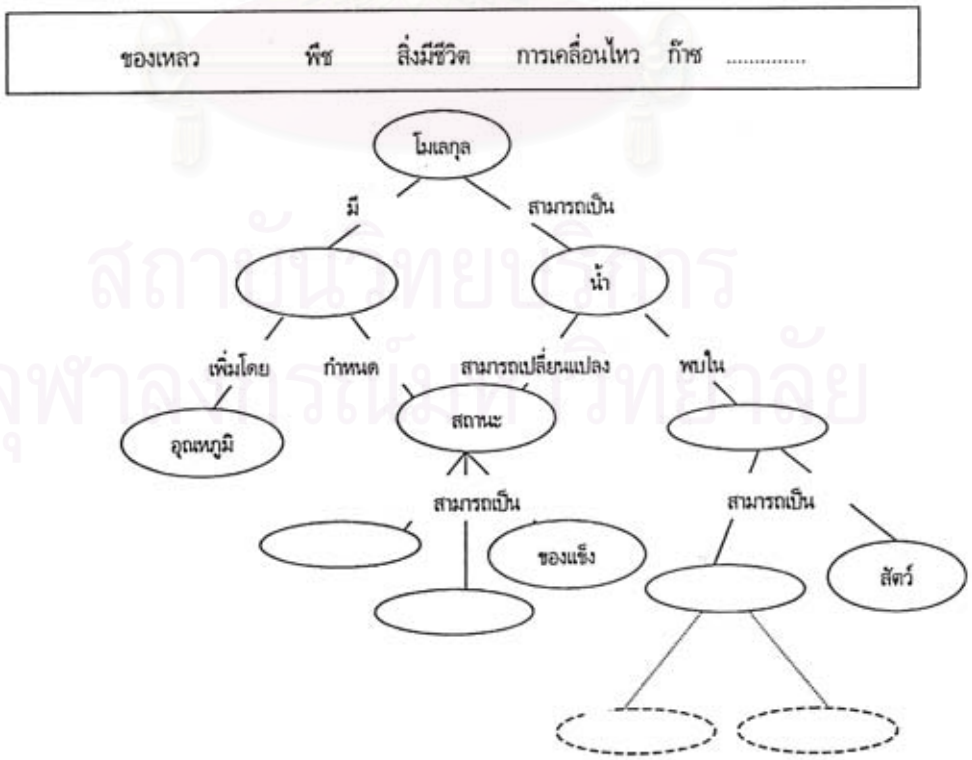
พืช	สิ่งมีชีวิต	หญ้า	วัว	สุนัข	ต้นมะม่วง	สัตว์

บาร์รูดีและบาร์เทล (Baroody and Bartels, 2000) ได้แบ่งประเภทของแผนผังออกเป็น 4 ประเภทดังนี้

1. แบบเติมมโนทัศน์ (fill-in task) เป็นการจัดแผนผังมโนทัศน์โดยมีการกำหนดมโนทัศน์ทั้งหมดมาให้ และมีการจัดแผนผังโครงสร้างที่ไม่สมบูรณ์มาให้ แล้วให้นักเรียนนำมโนทัศน์ที่กำหนดให้มาเติมใน แผนผังมโนทัศน์ให้สมบูรณ์ ดังตัวอย่าง



2. แบบเพิ่มมโนทัศน์ (add-on task) เป็นการจัดแผนผังมโนทัศน์โดยมีมโนทัศน์มาให้ และจัด แผนผังโครงสร้างที่ยังไม่สมบูรณ์มาให้ แล้วให้นักเรียนนำมโนทัศน์กำหนดมาเติมลงในช่องว่างของแผนผังโครงสร้าง โดยนักเรียนสามารถคิดมโนทัศน์ขึ้นมาเพิ่มได้ตามความเข้าใจและเพื่อให้แผนผังมโนทัศน์มี ความสมบูรณ์ยิ่งขึ้น ดังตัวอย่าง



3. แบบปลายปิด (close-ended list task) เป็นการจัดแผนผังมโนทัศน์โดยกำหนดมโนทัศน์มาให้ นักเรียนจัดแผนผังมโนทัศน์โดยใช้มโนทัศน์ทั้งหมดที่กำหนดมาให้และเขียนความสัมพันธ์ระหว่างมโนทัศน์ซึ่งทำให้การจัดแผนผังมโนทัศน์สมบูรณ์

ของเหลว	พืช	สิ่งมีชีวิต	การเคลื่อนไหว	ก๊าซ	สถานะ
อุณหภูมิ	ของแข็ง	ของเหลว	สัตว์	น้ำ	

4. แบบปลายเปิด (open-ended list task) เป็นการจัดแผนผังมโนทัศน์โดยกำหนดมโนทัศน์บางส่วนมาให้ โดยใช้คำมโนทัศน์ที่กำหนดให้และสามารถเพิ่มมโนทัศน์และเขียนความสัมพันธ์ระหว่างมโนทัศน์ ซึ่งทำให้แผนผังมโนทัศน์นั้นสมบูรณ์ยิ่งขึ้น

ของเหลว	พืช	สิ่งมีชีวิต	การเคลื่อนไหว	ก๊าซ	น้ำ
สถานะ	อุณหภูมิ	ของแข็ง	ของเหลว	สัตว์	

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การให้คะแนนแผนผังมโนทัศน์

ในการสร้างเกณฑ์การตรวจให้คะแนนแผนผังมโนทัศน์มีนักการศึกษาและนักวิจัยหลายท่าน ได้สร้างแนวทางไว้โดยมีรายละเอียดดังนี้

โนแวคและโกวิน (Novak & Gowin, 1984) ได้กำหนดเกณฑ์การให้คะแนนสำหรับแผนผังมโนทัศน์ไว้ดังนี้

1. ประพจน์ (propositions) ที่แสดงความสัมพันธ์ระหว่าง 2 มโนทัศน์ที่เชื่อมกันโดยคำเชื่อมมีความหมายที่สมเหตุผลและตรงกับความสัมพันธ์หรือไม่ ถ้ามีความหมายแต่ละข้อความที่แสดงความสัมพันธ์จะได้คะแนน 1 คะแนน

2. การจัดลำดับขั้นตอน (hierarchy) ในแผนผังมโนทัศน์มีการจัดลำดับหรือไม่ แต่ละมโนทัศน์ที่อยู่รองลงมาเป็นมโนทัศน์ที่เฉพาะเจาะจง และมีความกว้างขวางน้อยกว่ามโนทัศน์ที่อยู่ลำดับแรก ๆ ให้คะแนน 5 คะแนนของทุกระดับที่จัดลำดับได้สมเหตุสมผล

3. การเชื่อมระหว่างมโนทัศน์ (cross link) มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์หรือไม่ ถ้ามีการเชื่อมโยงอย่างสมเหตุสมผลและมีนัยสำคัญ ให้คะแนน 10 คะแนน และให้คะแนน 2 คะแนน ถ้าเชื่อมคำสมเหตุสมผลแต่ไม่ได้แสดงการสังเคราะห์ระหว่างชุดของมโนทัศน์หรือข้อความที่เกี่ยวข้องกัน การเชื่อมโยงระหว่างสายของมโนทัศน์สามารถชี้ให้เห็นความสามารถในการคิดสร้างสรรค์และต้องระมัดระวังเป็นพิเศษในการให้คะแนน การให้คำเชื่อมที่เกิดความคิดสร้างสรรค์ควรจะได้คะแนนพิเศษ

4. ตัวอย่าง (example) นักเรียนยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ที่สร้างขึ้นได้ถูกต้องจะได้คะแนนตัวอย่าง 1 คะแนน

บอล์ท (Bolte, 1999) ได้เสนอเกณฑ์การให้คะแนนแบบฮอลิสติก (holistic scoring) โดยแต่ละแผนผังมโนทัศน์มีคะแนนเต็ม 10 คะแนน โดยให้คะแนนของการจัดแผนผังมโนทัศน์ (organization) 6 คะแนน ความถูกต้อง (accuracy) 4 คะแนน ดังนี้

การจัดแผนผังมโนทัศน์ (organization) เป็นการตรวจสอบการจัดกลุ่มมโนทัศน์และการเชื่อมโยงให้คะแนนคือ

6 คะแนน แสดงได้ถึงความสมบูรณ์ของแผนผัง มีการเชื่อมโยงมโนทัศน์ที่สื่อความหมายถึงความสัมพันธ์ระหว่างมโนทัศน์ได้ลึกซึ้ง สร้างความสัมพันธ์ระหว่างกลุ่มของมโนทัศน์ได้อย่างชัดเจน ใช้คำเชื่อมได้อย่างเหมาะสม และเพิ่มมโนทัศน์ที่สัมพันธ์กับแผนผังเข้าไปได้อย่างถูกต้อง

5 คะแนน แสดงความเข้าใจของการเชื่อมโยงได้ลึกซึ้ง สร้างความสัมพันธ์ระหว่าง มโนทัศน์ได้ชัดเจน ใช้มโนทัศน์ที่กำหนดให้ทุกมโนทัศน์

4 คะแนน แสดงถึงความเข้าใจของการเชื่อมโยงได้แต่ไม่ลึกซึ้ง สร้างความสัมพันธ์ระหว่างมโนทัศน์ได้ไม่ค่อยชัดเจน มีบางมโนทัศน์ที่กำหนดให้แต่ไม่มีในแผนผังมโนทัศน์

3 คะแนน แสดงถึงความเข้าใจเพียงบางส่วนเท่านั้น เชื่อมความสัมพันธ์ระหว่างมโนทัศน์ไม่ชัดเจน มีมโนทัศน์หลายมโนทัศน์ที่กำหนดให้แต่ไม่มีในแผนผังมโนทัศน์

2 คะแนน แสดงถึงความเข้าใจได้เล็กน้อย เชื่อมความสัมพันธ์ระหว่างมโนทัศน์ไม่ชัดเจนมาก จัดกลุ่มมโนทัศน์ไม่ถูกต้อง ขาดมโนทัศน์หลักหลายมโนทัศน์

1 คะแนน แสดงถึงความเข้าใจได้น้อยมาก ใช้คำเชื่อมไม่ถูกต้อง ไม่แสดงถึงความสัมพันธ์ระหว่างมโนทัศน์ ขาดมโนทัศน์หลักหลายคำ

0 คะแนน ไม่แสดงถึงความเข้าใจ แสดงความสัมพันธ์ไม่ได้เลย ใช้คำเชื่อมไม่ถูกต้องเลย

ความถูกต้อง (accuracy) เป็นการตรวจสอบความชัดเจนและความผิดพลาดหรือความเข้าใจมโนทัศน์ที่คลาดเคลื่อน ให้คะแนนดังนี้

4 คะแนน ไม่ผิดพลาดเลย

3 คะแนน ผิดในรายละเอียดเล็กน้อย

2 คะแนน คลาดเคลื่อนเล็กน้อยและเป็นการคลาดเคลื่อนในมโนทัศน์รอง

1 คะแนน คลาดเคลื่อนในมโนทัศน์หลัก

0 คะแนน ไม่ถูกต้องเลย หรือคลาดเคลื่อนในมโนทัศน์หลักหลายมโนทัศน์

รุซ-พรีโม และคณะ (Ruiz-Primo et al, 2001) ได้ตรวจให้คะแนนแผนผังมโนทัศน์โดยการคำนวณให้คะแนน 3 ประเภท ดังนี้

1. proposition accuracy score คือ คะแนนความถูกต้องของคำมโนทัศน์ หรือผลรวมของคะแนนการเชื่อมคำมโนทัศน์ที่ถูกต้องที่นักเรียนทำได้ในแผนผังมโนทัศน์ที่สร้างขึ้น

2. convergence score คือ สัดส่วนของคะแนนรวมในการเชื่อมคำมโนทัศน์ที่ถูกต้องที่นักเรียนสามารถทำได้กับคะแนนเต็มของการเชื่อมคำมโนทัศน์ทั้งหมดในผังเกณฑ์

3. salience score คือ สัดส่วนของคะแนนรวมในการเชื่อมคำมโนทัศน์ที่ถูกต้องที่นักเรียนสามารถทำได้กับจำนวนการเชื่อมคำมโนทัศน์ทั้งหมดที่อยู่ในแผนผังของนักเรียน

เคน และคณะ (Klein et al, 2002) ได้เสนอเกณฑ์การให้คะแนนแผนผังมโนทัศน์ 3 ประเภท ดังนี้

1. original คือ คะแนนความถูกต้องของแผนผังมโนทัศน์เมื่อเทียบกับแผนผังมโนทัศน์ที่ใช้เป็นเกณฑ์

2. supplementary credit คือ ให้คะแนนความน่าเชื่อถือสำหรับการเชื่อมมโนทัศน์สำคัญถูกต้อง เหมือนกับแผนผังมโนทัศน์ที่ใช้เป็นเกณฑ์

3. quality rating คือ การให้คะแนนความเป็นเหตุเป็นผลโดยมีการจัดแบ่งเป็น 4 กลุ่ม คือ (0) illogical คือ ปราศจากเหตุผล (1) pragmatic คือ ความเป็นเหตุเป็นผล (2) scientific คือ หลักเกณฑ์ทางวิทยาศาสตร์ (3) principled คือ มีหลักการ

การตรวจให้คะแนนของแผนผังมโนทัศน์มีหลายวิธี ซึ่งในการนำไปใช้ขึ้นอยู่กับรูปแบบของแผนผังมโนทัศน์ที่ใช้ โดยผู้ประเมินสามารถประยุกต์หรือสร้างเกณฑ์การให้คะแนนได้ด้วยตนเอง

ประโยชน์ของแผนผังมโนทัศน์

การนำแผนผังมโนทัศน์ไปใช้ในการศึกษาอย่างกว้างขวาง นักการศึกษาหลายท่านได้กล่าวถึงการนำแผนผังมโนทัศน์ไปใช้ประโยชน์ ดังนี้

มนัส บุญประกอบ (2542) กล่าวว่าการนำแผนผังมโนทัศน์ไปใช้สามารถใช้ได้อย่างกว้างขวางหลายประการ เช่น

1. การระดมสมอง เหมาะแก่การใช้แผนผังมโนทัศน์ประเภทกระจายออกอย่างยิ่ง เพราะเมื่อคิดถึงเรื่องใดเรื่องหนึ่งให้เขียนเรื่องนั้นลงไปก่อน แล้วเขียนเส้นล้อมกรอบไว้ จากนั้นให้คิดต่อว่ามีอะไรบ้างที่เกี่ยวข้อง คิดได้เท่าไรเขียนลงไปทั้งหมดจึงเขียนเส้นโยงไปยังเรื่องหลัก เมื่อได้มาก ๆ ขึ้นก็อาจจำเป็นต้องการจัดกลุ่มในภายหลัง การเขียนเช่นนี้เป็นลักษณะธรรมชาติและเป็นอิสระมา ใช้ได้ทั้งการระดมสมองคนเดียว หรือเป็นกลุ่มย่อย และกลุ่มใหญ่

2. การวางแผนงาน เช่น การเตรียมตัวสอบแข่งขัน การเตรียมทำกิจกรรมชุมนุม การเตรียมฝึกนักกีฬา การวางแผนการวิจัย เป็นต้น การเขียนถ่ายทอดความคิดออกมาเป็นแผนผังมโนทัศน์ช่วยให้มองเป็นรูปธรรมชัดเจนภายในกระดาษเพียงหน้าเดียว ซึ่งสามารถใช้ประโยชน์ในการสื่อสารความเข้าใจแก่กัน ดุจแผ่นพิมพ์เขียวของสถาปนิกทีเดียว

3. การออกแบบ การออกแบบพัฒนาหลักสูตร การออกแบบการเชื่อมโยงข้อสนเทศบน โฮมเพจของตนเอง ลักษณะของแผนผังมโนทัศน์เอื้อประโยชน์ในแง่นี้ได้อย่างดี เพราะแผนผังมโนทัศน์เน้นถึงความสัมพันธ์ระหว่างจุดต่าง ๆ มากมายได้เสมอทุกรูปแบบ

4. การเขียนเรื่องราว ก่อนที่จะเขียนบทความ เรียงความ รายงาน หรือเอกสาร ตำรา ต้องคิดโครงเรื่องหลักเพื่อแจกแจงไปสู่รายละเอียดในภายหลัง ลักษณะพิเศษของแผนผังมโนทัศน์ที่แสดงความสัมพันธ์ได้หลายรูปแบบดังกล่าวมา แล้วจะช่วยในเชิงความสัมพันธ์ของภาพรวมและส่วนขยายไปจนถึงรายละเอียดที่ต้องการและมีความซับซ้อนมาก ๆ ได้

5. เครื่องมือการเรียนรู้ การเขียนสรุปเรื่องราวแต่ละบทแต่ละตอนด้วยแผนผังมโนทัศน์ อาจช่วยทำให้จดจำได้เป็นอย่างดี เพียงใช้คำศัพท์เฉพาะต่าง ๆ มาจัดเรียงความสัมพันธ์

6. เครื่องมือการประเมิน ครูอาจใช้แผนผังมโนทัศน์ ประเมินผู้เรียนแทนการเติมคำลงในช่องว่างหรือแทนการเขียนตอบแบบอัตนัย เช่น ในเบื้องต้นของแบบฝึกหัดท้ายบทเรียน ท่านอาจเขียนเป็นแผนผังมโนทัศน์ที่เว้นว่างเป็นบางกรอบสำหรับให้เติมข้อความ กรณีที่ผู้เรียนเขียนแผนผังมโนทัศน์ด้วยตนเองแล้ว ครูอาจถามให้ตอบอย่างอิสระและเปิดกว้างโดยกำหนดให้เขียนเป็นแผนผังรูปแบบใด ๆ โดยมีเกณฑ์การตรวจให้คะแนนที่คำนึงถึงความถูกต้องแง่เนื้อหา หลักการ จำนวนคำมโนทัศน์ หรือคำศัพท์เฉพาะที่ใช้ระดับความซับซ้อนหรือชนิดของแผนผัง และความแปลกใหม่ในเชิงความคิดสร้างสรรค์

7. เครื่องมือในการนำเสนอเรื่องราว หรือการถ่ายทอดความรู้

8. การทำวิจัยเพื่อขยายองค์ความรู้ หรือเพื่อทำผลงานขอเลื่อนตำแหน่งวิชาการ

แลนซิง (Lanzing, 2000) ได้กล่าวถึงประโยชน์ของการใช้แผนผังมโนทัศน์ ดังนี้

1. การสร้างแนวคิด เช่น การระดมสมอง

2. การออกแบบสิ่งที่มีโครงสร้างสลับซับซ้อน เช่น สื่อขนาดใหญ่ (hypermedia) ตำรา (long text) เว็บไซต์ขนาดใหญ่ (large web sites)

3. การสื่อสารในแนวคิดที่ซับซ้อน

4. การช่วยให้เกิดการเรียนรู้ด้วยการผสานองค์ความรู้ที่มีอยู่และองค์ความรู้ใหม่เข้าด้วยกัน โดยแสดงให้เห็นชัดเจน

5. การช่วยประเมินความเข้าใจที่ถูกต้องหรือวินิจฉัยความเข้าใจที่คลาดเคลื่อน

แมคเคบ (McCabe, 1995) ได้กล่าวถึงประโยชน์ของแผนผังมโนทัศน์ที่นำไปใช้ในการเรียนการสอน ดังนี้

1. เป็นเครื่องมือสำหรับนักเรียนนักศึกษา เพื่อ

1.1 เพื่อจัดการกับเนื้อหาสาระในวิชาต่าง ๆ

1.2 เพื่อเขียนแผนผังโครงสร้างองค์ความรู้ของวิชาต่าง ๆ ที่เรียนรู้

1.3 เพื่อสอดแทรกมโนทัศน์ใหม่ ๆ เข้าไปในโครงสร้างองค์ความรู้ที่มีอยู่เดิม

1.4 เพื่อเตือนความจำในสิ่งที่ต้องใช้เวลายาวนาน

1.5 เพื่อปรับปรุงการสอนอย่างมีประสิทธิภาพ

2. เป็นเครื่องมือในการวางแผนงานและการสอนของครู

2.1 เพื่อประเมินความรู้ของนักเรียนที่มีอยู่อย่างรวดเร็ว

2.2 เพื่อวางแผนและปรับปรุงกิจกรรมการเรียนการสอน

2.3 เพื่อประเมินการเรียนรู้ของนักเรียน

2.4 เพื่อทบทวนประเด็นต่าง ๆ

2.5 สำหรับสำรวจกิจกรรมการสอนที่เฉพาะเจาะจง

บาร์รูดีและบาร์เทล (Baroody and Bartels, 2001) ได้กล่าวว่าแผนผังมโนทัศน์เป็นเครื่องมือประเมินกระบวนการคิดประเภทหนึ่ง และได้กล่าวถึงมีประโยชน์ของการใช้แผนผังมโนทัศน์ในการประเมินไว้ดังนี้

1. แผนผังมโนทัศน์สามารถใช้ในการวินิจฉัยและเป็นข้อมูลย้อนกลับในการจัดการเรียนการสอนได้เช่นเดียวกับการตัดเกรด นอกจากนี้การสร้างแผนผังมโนทัศน์ของนักเรียนสามารถทำให้ครูได้ทราบถึงระดับความเข้าใจในความคิดรวบยอด และสารสนเทศที่ได้สามารถไปใช้ในการวางแผนหรือปรับปรุงการเรียนการสอนได้

2. ใช้เป็นเครื่องมือในการจัดการเรียนการสอน การสังเกตพฤติกรรมการสร้างแผนผังมโนทัศน์ทำให้ทราบจุดบกพร่องของนักเรียน และเป็นเครื่องมือที่ทำให้ทราบการเรียนรู้ของผู้เรียนได้ชัดเจนยิ่งขึ้น

3. แผนผังมโนทัศน์สามารถใช้ประเมินการทำงานกลุ่มและพัฒนาการรายบุคคลได้

4. แผนผังมโนทัศน์สามารถใช้แบบทดสอบทางด้านการเขียนและเป็นวิธีการประเมินทางเลือกใหม่สำหรับครู กิจกรรมการทดสอบโดยใช้แผนผังมโนทัศน์โดยการฟังการอภิปรายเกี่ยวกับการสร้างแผนผังทำให้ครูผู้สอนเข้าใจความคิดของนักเรียนได้ชัดเจนขึ้น

กำพล ดำรงค์วงศ์ (2540) ได้สรุปผลการใช้แผนผังมโนทัศน์ในการเรียนการสอนได้ดังนี้

1. การสร้างแผนผังมโนทัศน์เป็นเทคนิคการสอนที่สามารถนำมาใช้สอนในวิชาต่าง ๆ ได้หลายวิชา ได้แก่ วิชาชีววิทยา วิทยาศาสตร์ทั่วไป การอ่าน สิ่งแวดล้อม เศรษฐศาสตร์ สังคมวิทยา เคมี ฟิสิกส์ เป็นต้น ทำให้นักเรียนเข้าใจอย่างสมบูรณ์มากขึ้น ซึ่งเป็นเครื่องมือการเรียนรู้ได้เป็นอย่างดี

2. แผนผังมโนทัศน์เป็นเครื่องมือประเมินผลได้เป็นอย่างดี สามารถใช้วัดความรู้ความสามารถในการใช้เหตุผลของนักเรียน และวัดมโนทัศน์ที่คลาดเคลื่อนของนักเรียนได้ ทำให้ครูทราบโครงสร้างความรู้ของนักเรียน

3. นักเรียนที่ได้รับการฝึกฝนให้สร้างแผนผังมโนทัศน์ขึ้นได้ด้วยตนเองทำให้เกิดการเรียนรู้อย่างมีความหมาย เกิดความคงทนในการเรียนรู้

4. นักเรียนที่ได้รับการสอนโดยใช้แผนผังมโนทัศน์มีคะแนนผลสัมฤทธิ์ทางการเรียนสูงกว่านักเรียนที่ได้รับการสอนด้วยวิธีปกติ

5. นักเรียนมีเจตคติที่ดีต่อการเรียนโดยใช้แผนผังมโนทัศน์

6. นักเรียนที่ได้รับการสอนโดยใช้แผนผังมโนทัศน์ สามารถเปลี่ยนมโนทัศน์ที่คลาดเคลื่อนให้เป็นมโนทัศน์ที่ถูกต้องได้มากกว่าการสอนด้วยวิธีปกติ

จากการศึกษาสามารถสรุปประโยชน์ของการนำแผนผังมโนทัศน์ไปใช้ในการศึกษาได้ดังนี้

1. ใช้เป็นเครื่องมือในการเตรียมการสอนของครู ครูสามารถจัดกิจกรรมการเรียนการสอนโดยใช้แผนผังมโนทัศน์เพื่อสำรวจความรู้ของนักเรียนแล้วนำไปวางแผนการสอนให้เหมาะสมกับนักเรียนและใช้ในการจัดลำดับเนื้อหาสาระที่จะสอน

2. ใช้เป็นเครื่องมือในการประเมินผลการเรียนรู้และประเมินความคิดรวบยอดของนักเรียน โดยการให้นักเรียนสรุปสิ่งที่เรียนเป็นแผนผังมโนทัศน์หรือตอบข้อสอบโดยใช้แผนผังมโนทัศน์เพื่อแสดงความเข้าใจในการเขียนตอบและเพื่อวินิจฉัยความเข้าใจที่คลาดเคลื่อนของนักเรียน

3. ใช้เป็นเครื่องมือในการเรียนรู้สำหรับนักเรียน เพื่อให้เกิดการเรียนรู้อย่างมีความหมายโดยการใช้กรอบมโนทัศน์เพื่อสร้างความหมายจากสิ่งที่เรียน จะทำให้นักเรียนจดจำไปได้บ้างและมีความคงทนเพราะจะทำให้เห็นความสัมพันธ์ของสิ่งที่เรียนไปทั้งหมด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 3 การตรวจสอบคุณภาพของเครื่องมือ

ความตรง (Validity) เป็นคุณสมบัติที่สำคัญของเครื่องมือวัดผล ซึ่งเกี่ยวข้องกับคุณภาพด้านความถูกต้องของผลที่ได้จากการวัด ทำให้สามารถนำคะแนนที่ได้ไปแปลความหมาย ถึงสิ่งที่มุ่งวัดได้อย่างเหมาะสม (ศิริชัย กาญจนวาสี, 2544)

ตามทฤษฎีการทดสอบดั้งเดิม ได้ให้คำจำกัดความของคำว่า “ความตรง” ไว้ดังนี้

1) ความตรง หมายถึง ความถูกต้องแม่นยำของเครื่องมือในการวัดสิ่งที่ต้องการจะวัด

ความตรงเป็นความใกล้เคียงกันระหว่างค่าที่วัดได้กับค่าที่แท้จริงถ้าผลการวัดได้ค่าที่ใกล้เคียงกับค่าที่แท้จริงเพียงใด คือว่าการวัดมีความตรงมากขึ้นเพียงนั้น ดังนั้นความตรงจึงเป็นคุณสมบัติที่สำคัญที่สุดของเครื่องมือหรือแบบสอบ แบบสอบที่มีความตรงสูงจะเป็นเครื่องมือที่วัดคุณลักษณะที่ต้องการได้ตรงหรือใกล้เคียงกับค่าที่แท้จริงมาก คะแนนที่ได้จากแบบสอบที่มีความตรงสูง คะแนนนั้นสามารถบ่งบอกถึงสภาพที่แท้จริงของลักษณะที่มุ่งวัดนั้นได้เป็นอย่างดี

คำจำกัดความนี้ถือเป็นนิยามของความตรงในเชิงความคิดหรือทฤษฎีซึ่งไม่สามารถประมาณค่าได้โดยตรง เพราะไม่ทราบค่าที่แท้จริงของสิ่งที่มุ่งวัด จึงต้องหาวิธีประมาณค่าความตรงโดยกำหนดการประมาณค่าที่แท้จริงด้วยคำจำกัดความเชิงปฏิบัติการ ดังจะกล่าวต่อไปนี้

2) ความตรง หมายถึง “ความสอดคล้องหรือความเหมาะสมของผลการวัดกับเนื้อเรื่องหรือเกณฑ์ หรือทฤษฎีเกี่ยวกับลักษณะที่มุ่งวัด” หรืออาจนิยามเชิงปฏิบัติการในลักษณะของเทคนิคการคำนวณได้ว่า

ความตรง หมายถึง “ค่าสัมประสิทธิ์สหพันธ์ระหว่างคะแนนที่ได้จากเครื่องมือกับคะแนนเกณฑ์ภายนอกที่เป็นอิสระอื่น ๆ ซึ่งเชื่อถือได้ว่าสามารถวัดสิ่งที่ต้องการวัดนั้นได้

คำจำกัดความนี้ทำให้เราสามารถประมาณค่าความตรงของแบบสอบได้ แต่ต้องสามารถระบุ โครงสร้างของคุณลักษณะ เนื้อหา หรือเกณฑ์ภายนอกที่เป็นตัวแทนที่ต้องการวัดได้อย่างเหมาะสม เพื่อใช้เป็นเกณฑ์เทียบถึงความถูกต้องเหมาะสม หรือสอดคล้องของคะแนนจากแบบสอบในการวัดคุณลักษณะที่ต้องการนั้น

คำจำกัดความนี้ทำให้เราสามารถประมาณค่าความตรงของแบบสอบได้ แต่ต้องสามารถระบุ โครงสร้างของคุณลักษณะ เนื้อหา หรือเกณฑ์ภายนอกที่เป็นตัวแทนลักษณะที่ต้องการวัดได้อย่างเหมาะสม เพื่อใช้เป็นเกณฑ์เทียบถึงความถูกต้องเหมาะสม หรือความสอดคล้องของคะแนนจากแบบสอบในการวัด คุณลักษณะที่ต้องการนั้น

ประเภทและวิธีตรวจสอบความตรง (ศิริชัย กาญจนวาสี, 2544)

การตรวจสอบความตรงเป็น “กระบวนการรวบรวมและวิเคราะห์หลักฐานเพื่อการสนับสนุนความเหมาะสมและความถูกต้องของการนำคะแนนจากเครื่องมือวัดไปสรุปอ้างอิงถึงคุณลักษณะที่มุ่งวัด” การตรวจสอบความตรงของคะแนนจากแบบสอบถาม จะต้องมีความชัดเจนตั้งแต่ต้นว่า “ลักษณะที่ต้องการสรุปอ้างอิงไปถึงนั้นคืออะไร” เพื่อทำการรวบรวมหลักฐานอันเป็นประโยชน์ต่อการตัดสินความเหมาะสมและความถูกต้องได้ตรงตามเป้าหมายของการนำคะแนนนั้นไปใช้ ในการตรวจสอบความตรง สามารถจำแนกตามเป้าหมายสำคัญได้ 3 ประเภท ได้แก่

1) การตรวจสอบความตรงตามเนื้อเรื่อง (content validation) เป็นการสรุปอ้างอิงถึงมวลเนื้อเรื่อง ความรู้หรือประสบการณ์ (domain of contern) ที่แบบสอบมุ่งวัดค่าการวัดได้ผลครอบคลุมและเป็นตัวแทนมวลความรู้หรือประสบการณ์นั้นดีเพียงไร

2) การตรวจสอบความตรงตามเกณฑ์สัมพันธ์ (criterion-related validation) เป็นการสรุปอ้างอิงถึงสมรรถนะการดำเนินงาน (practical performance) ของสิ่งที่มุ่งวัดว่า การวัดได้ผลสอดคล้องกับผลการดำเนินงานนั้นเพียงไร ซึ่งเกณฑ์ที่ใช้เทียบความสัมพันธ์อาจเป็นสถานภาพการดำเนินงานที่เป็นอยู่จริงใน ปัจจุบันหรืออาจเป็นผลสำเร็จของการปฏิบัติงานนั้นในอนาคต

3) การตรวจสอบความตรงเชิงทฤษฎี (construct validation) เป็นการสรุปอ้างอิงถึงสภาวะ หรือ โครงสร้างทางจิตวิทยา (psychological construct) ของลักษณะที่มุ่งวัดนั้นว่า การวัดได้ผลตรงตามทฤษฎีของลักษณะทางจิตวิทยานั้นดีเพียงไร

การตรวจสอบความตรงต่างประเภทกัน ย่อมให้หลักฐานการสรุปอ้างอิงของคะแนนไปยังเป้าหมายของสิ่งที่มุ่งวัดต่างกัน และไม่สามารถใช้ทดแทนกันได้ แต่นำมาใช้สนับสนุนซึ่งกันและกันได้ หลักฐานที่ใช้สนับสนุนความตรงของแบบสอบควรใช้หลักฐานจากการตรวจสอบความตรงมากกว่า 1 ประเภท วิธีการตรวจสอบความตรงแต่ละประเภท มีรายละเอียดดังต่อไปนี้

การตรวจสอบความตรงตามเนื้อเรื่อง (Content Validity) (ศิริชัย กาญจนวาสี, 2544)

ความตรงตามเนื้อเรื่องเป็นความตรงพื้นฐานที่มีความสำคัญต่อมาพัฒนาแบบสอบประเภทต่าง ๆ ทางจิตวิทยาและการศึกษา โดยหลักการของการวิเคราะห์ความตรงตามเนื้อเรื่องมีเป้าหมายในการตรวจสอบว่า “ข้อสอบที่ใช้มีความเพียงพอและเป็นตัวแทนมวลความรู้ / ประสบการณ์ (domain of content) หรือ แนวคิดเชิงทฤษฎี (theoretical concept) ที่มุ่งวัดเพียงไร” ความตรงตามเนื้อเรื่อง จึงขึ้นกับผลเชิงประจักษ์ ที่ได้จากการวัดว่าสามารถใช้เป็นตัวแทนบ่งบอกมวลเนื้อเรื่อง / ประสบการณ์เฉพาะที่ต้องการวัดได้ดีเพียงไร

ข้อควรพิจารณาในการตรวจสอบความตรงตามเนื้อเรื่อง (ศิริชัย กาญจนวาสี, 2544)

การตรวจสอบความตรงตามเนื้อเรื่องมีข้อควรพิจารณาที่สำคัญดังนี้

1) การให้น้ำหนักความสำคัญของจุดมุ่งหมายการเรียนรู้

การกำหนดเป้าหมายของการเรียนรู้โดยทั่วไป บางครั้งมักใช้ความสะดวกในการให้น้ำหนักความสำคัญที่เท่ากันสำหรับแต่ละจุดมุ่งหมาย แต่ตามสภาพความเป็นจริงควรกำหนดความสำคัญด้วยการจัดเรียงอันดับความสำคัญของแต่ละจุดมุ่งหมาย หรืออาจกำหนดน้ำหนักความสำคัญด้วยมาตรประมาณค่า 5 ช่วง เช่น 5 = สำคัญมากที่สุด 4 สำคัญมาก 3 = สำคัญปานกลาง 2 = สำคัญน้อย 1 = สำคัญน้อยมาก เป็นต้น เมื่อได้น้ำหนักความสำคัญแล้ว ควรให้ผู้เชี่ยวชาญตัดสินว่าจำนวนข้อสอบที่เหมาะสม / และเป็นตัวแทนมวลเนื้อเรื่อง / ประสบการณ์ที่มุ่งวัดนั้น

2) มิติที่ต้องการให้ผู้เชี่ยวชาญตรวจสอบ

การระบุมิติและทิศทางในการตัดสินที่ชัดเจนให้แก่กลุ่มผู้เชี่ยวชาญว่า ต้องการให้ตัดสินอะไรบ้าง ผลการตัดสินควรมีทิศทางอย่างไร เช่น ให้ตัดสินความสอดคล้องระหว่างสิ่งที่ข้อสอบมุ่งวัดกับจุด มุ่งหมาย / หรือเนื้อเรื่อง ในบางสถานการณ์อาจจำเป็นต้องให้ผู้เชี่ยวชาญช่วยพิจารณาความเหมาะสมของ คำจำกัดความ ความชัดเจนของภาษาที่ใช้ ระดับพฤติกรรมที่มุ่งวัด ขั้นการคิดแก้ปัญหา รูปแบบข้อสอบ และวิธีตอบที่เหมาะสม เป็นต้น

3) การตรวจสอบความสอดคล้องระหว่างข้อสอบกับจุดมุ่งหมาย

ควรกำหนดเงื่อนไขให้ผู้เชี่ยวชาญมีการตัดสินอย่างเป็นระบบ เช่น ควรเขียนข้อสอบแยกเป็นรายข้อในบัตรข้อสอบ หรือจัดทำเป็นช่องตารางที่สามารถให้คำวิจารณ์ได้เป็นรายข้อ เพื่อให้ผู้เชี่ยวชาญตรวจสอบความถูกต้องของคำตอบ ตัดสินว่าข้อสอบวัดจุดมุ่งหมายใด และควรมีการบันทึกผลการตัดสินอย่างเป็นรูปแบบที่สามารถนำมาวิเคราะห์ผลการตัดสินได้อย่างชัดเจน ส่วนระดับของความสอดคล้องอาจใช้รูปแบบมาตรประมาณค่า 5 สเกล หรืออาจจัดระดับเป็นสอดคล้อง (+1) ไม่แน่ใจ (0) และไม่สอดคล้อง (-1)

4) วิธีการสรุปผลการตัดสิน ความตรงตามเนื้อเรื่อง

การตรวจสอบความครอบคลุมและความเป็นตัวแทนของมวลผลสัมฤทธิ์ที่มุ่งวัดดูเหมือนจะเป็นการตรวจสอบในเชิงคุณภาพมากกว่าในเชิงปริมาณ ซึ่งจะต้องเน้นความสำคัญของการตัดสินตามการวิเคราะห์เนื้อเรื่อง อย่างไรก็ตามมีดัชนีบ่งชี้ในเชิงประมาณหลายลักษณะที่ควรเลือกใช้ให้เหมาะสม เช่น ร้อยละของจำนวนข้อสอบที่สอดคล้องกับจุดมุ่งหมาย ร้อยละของจำนวนข้อสอบที่สอดคล้องระดับสูงกับ จุดมุ่งหมาย สัมประสิทธิ์สหสัมพันธ์ระหว่างน้ำหนักความสำคัญของจุดมุ่งหมายกับจำนวนข้อสอบที่ใช้วัด จุดมุ่งหมาย ร้อยละของจุดมุ่งหมายที่ไม่ได้วัดโดยข้อสอบใดเลย เป็นต้น ดัชนีแต่ละตัวอย่างมีหลักการใช้และความหมายต่างกัน การเลือกใช้

ดัชนีต่างกันอาจนำไปสู่ข้อสรุปของความตรงตามเนื้อเรื่องต่างกันได้ จึงต้องเลือกใช้และแปลผลด้วยความสมเหตุสมผล

การตรวจสอบความตรงตามเกณฑ์สัมพันธ์ (Criterion- related Validity)

(ศิริชัย กาญจนวาสี, 2544)

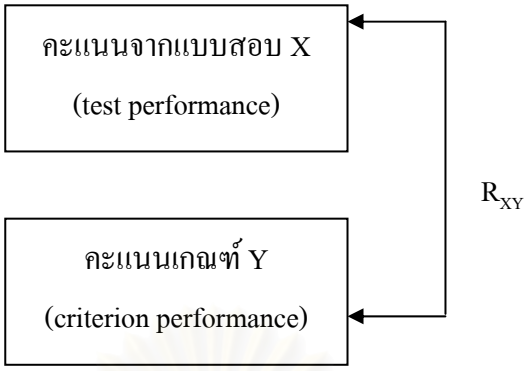
ความตรงตามเกณฑ์สัมพันธ์เป็นคุณสมบัติด้านความสอดคล้องสัมพันธ์กันระหว่างคะแนนจากแบบสอบกับ เกณฑ์ ภายนอก (criterion) ที่สามารถใช้วัดลักษณะที่ต้องการนั้นได้ เกณฑ์ภายนอกอาจเป็นสมรรถนะ ผลการปฏิบัติ พฤติกรรม หรือผลการดำเนินงานที่ปรากฏของคุณลักษณะนั้น เช่น ในสถานการณ์ของการสร้างแบบสอบคัดเลือกนักเรียนเข้าศึกษาต่อในสถาบันการศึกษา หรือการสร้างแบบสอบคัดเลือกผู้สมัครเข้าทำงานในบริษัท ก่อนที่จะนำคะแนนจากแบบสอบไปตัดสินใจเกี่ยวกับการคัดเลือกได้อย่างถูกต้องและเหมาะสม คงจะต้องรู้ถึงคุณภาพของคะแนนสอบที่ได้ว่าสามารถนำไปสรุปอ้างอิงถึงความสามารถทางด้านการเรียนวิชาการในสถาบันการศึกษา หรือความสามารถในการทำงานด้านนั้นได้เป็นอย่างดีเสียก่อนหลักฐานที่แสดงระดับความสัมพันธ์ระหว่างคะแนนสอบกับเกณฑ์ที่เป็นสมรรถนะ พฤติกรรม หรือความสามารถที่ปรากฏนี้เรียกว่า ความตรงตามเกณฑ์สัมพันธ์

ประเภทของความตรงตามเกณฑ์สัมพันธ์

ความตรงตามเกณฑ์สัมพันธ์นี้ สามารถจำแนกได้เป็น 2 ประเภท ตามลักษณะของเกณฑ์ที่นำมาใช้หาความสัมพันธ์กับคะแนนสอบ ดังนี้

1) ความตรงตามสภาพ หรือความตรงร่วมสมัย (concurrent Validity)

ความตรงตามสภาพ หรือ ความตรงร่วมสมัย เป็นความตรงตามเกณฑ์ซึ่งสัมพันธ์กับ สมรรถนะการดำเนินงานในสภาพปัจจุบัน (current perfomance) การประมาณค่าความตรงตามสภาพของแบบสอบให้ความสนใจประมาณ สถานภาพปัจจุบัน (present status) ของสมรรถนะที่มุ่งวัด โดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง คะแนนจากแบบสอบ กับ คะแนนเกณฑ์จากเครื่องมืออื่นที่สามารถใช้บ่งบอกสถานภาพปัจจุบันของลักษณะที่มุ่งวัดนั้นได้ เนื่องจากเครื่องมือทั้งสองทำการวัดในเวลาเดียวกัน สัมประสิทธิ์สหสัมพันธ์ในทางบวกที่สูงแสดงถึงคะแนนจากแบบสอบถามสามารถใช้เป็นตัวบ่งชี้ที่ดีของสถานภาพของลักษณะที่มุ่งวัดนั้น ดังแผนภาพที่ 3



แผนภาพที่ 2.3 แนวคิดของการตรวจสอบความตรงตามสภาพ (ศิริชัย กาญจนวาสี, 2544)

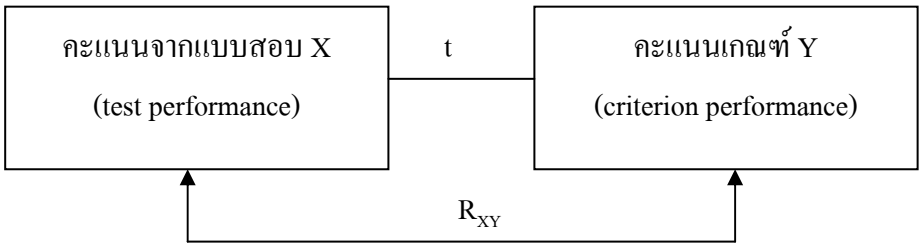
จากสูตร

$$R_{XY} = \frac{n\sum xy - \sum x \sum y}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

R_{xy} = สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน x และคะแนน y

2) ความตรงเชิงทำนาย (Predictive Validity)

ความตรงเชิงทำนายเป็นความตรงตามเกณฑ์ซึ่งสัมพันธ์กับ สมรรถนะการดำเนินงานในอนาคต (future performance) การประมาณค่าความตรงเชิงทำนายของแบบสอบให้ความสนใจประมาณ สถานภาพอนาคต (future status) ของคุณลักษณะที่มุ่งวัดโดยการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง คะแนนจากแบบสอบถาม กับคะแนนเกณฑ์ จากเครื่องมือที่สามารถใช้บ่งบอกสำเร็จ ของลักษณะที่มุ่งวัดนั้นในอนาคต เนื่องจากเครื่องมือทั้งสองทำการวัดในเวลาต่างกัน โดยแบบสอบที่สร้างทำการวัดเพื่อให้ได้คะแนนสอบในปัจจุบัน แต่อีกเครื่องมือหนึ่งต้องทิ้งช่วงเวลาทำการวัดในเวลาต่อมาเพื่อให้ได้คะแนนเกณฑ์ในอนาคต สำหรับนำมาใช้คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ ดังภาพที่ 4



ภาพที่ 2.4 แนวคิดของการตรวจสอบความตรงเชิงทำนาย (ศิริชัย กาญจนวาสี, 2544)

ข้อควรพิจารณาในการประมาณค่าความตรงตามเกณฑ์สัมพันธ์

จุดสำคัญของการตรวจสอบความตรงตามเกณฑ์สัมพันธ์ได้แก่ การคัดเลือกตัวแปรเกณฑ์ที่มีความเหมาะสมในด้านของการเป็นตัวแปรเกณฑ์ที่ตรงประเด็น และเป็นตัวแปรเกณฑ์ที่มีความตรง เนื่องจากเกณฑ์ที่ใช้จะมีผลกระทบโดยตรงต่อคุณภาพของแบบสอบหรือเครื่องมือที่ต้องการตรวจสอบ

1) ตัวแปรเกณฑ์ที่ตรงประเด็น (relevant criterion)

ควรคัดเลือกตัวแปรเกณฑ์ภายนอกที่มุ่งวัดคุณลักษณะที่สัมพันธ์หรือสอดคล้องตรงประเด็นกับลักษณะที่มุ่งวัดโดยแบบสอบซึ่งสร้างขึ้นมา กล่าวคือ แบบสอบและเครื่องมือที่ใช้เป็นเกณฑ์เปรียบเทียบความสัมพันธ์จะต้องมุ่งวัดลักษณะที่ตรงประเด็นกัน หรือเกี่ยวข้องกัน

2) ตัวแปรเกณฑ์ที่มีความตรง (valid criterion)

ควรคัดเลือกตัวแปรเกณฑ์ภายนอกที่เชื่อถือได้ว่าสามารถวัดได้ตรงตามลักษณะที่มุ่งวัดนั้น กล่าวคือเครื่องมือที่ใช้เป็นเกณฑ์เปรียบเทียบความสัมพันธ์จะต้องมีหลักฐานยืนยันได้ว่าเครื่องมือนั้นมีความตรง

การตรวจสอบความตรงเชิงทฤษฎี (Construct Validity) (ศิริชัย กาญจนวาสี, 2544)

ความตรงตามเนื้อเรื่องและความตรงตามเกณฑ์สัมพันธ์เป็นคุณสมบัติของแบบสอบที่นิยมตรวจสอบกันอย่างกว้างขวาง เพื่อให้ทราบถึงความถูกต้องและเหมาะสมในการนำคะแนนสอบไปใช้ตามจุดมุ่งหมายเฉพาะอย่าง ความตรงตามเนื้อเรื่องเป็นหลักฐานที่ช่วยบ่งบอกถึงความเหมาะสมของคะแนนสอบในด้านความครอบคลุมและเป็นเนื้อตัวแทนมวลเนื้อเรื่อง /ประสบการณ์ที่มุ่งวัดซึ่งเป็นประโยชน์มากสำหรับพัฒนาและตรวจสอบคุณภาพของแบบสอบผลสัมฤทธิ์ ส่วน ความตรงตามเกณฑ์สัมพันธ์ เป็นหลักฐานแสดงถึงความเหมาะสมของคะแนนสอบในด้านการบ่งชี้สภาพ

หรือทำนายผลการดำเนินงานตามเกณฑ์ที่ต้องการ แต่ในขณะเดียวกับความตรงทั้ง 2 ประเภทนี้ต่างมีข้อจำกัดที่สำคัญหลายประการดังกล่าวมาแล้ว

แผนภาพที่ 2.5 มโนทัศน์ของการตรวจสอบความตรงเชิงทฤษฎี (ศิริชัย กาญจนวาสี, 2544)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิธีการตรวจสอบความตรงเชิงทฤษฎี (ศิริชัย กาญจนวาสี, 2544)

การตรวจสอบความตรงเชิงทฤษฎี จำเป็นต้องอาศัยการสรุปอ้างอิงถึงลักษณะที่มุ่งวัด โดยใช้หลักฐานสนับสนุนหลายอย่าง ความตรงตามเนื้อเรื่องและความตรงตามเกณฑ์สัมพันธ์ ถือว่าเป็นหลักฐานส่วนหนึ่ง สำหรับสนับสนุนความตรงเชิงทฤษฎี ลำพังเพียงค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสอบกับคะแนนเกณฑ์ยังไม่เพียงพอสำหรับสรุปผลความตรงเชิงทฤษฎี จำเป็นต้องอาศัยหลักฐานความสอดคล้องกับทฤษฎี เช่น หลักฐานการตรวจสอบผลการวัดที่เป็นไปตามคำทำนายของทฤษฎี หลักฐานการเปรียบเทียบความสัมพันธ์กับผลการวัดด้วยเครื่องมือที่มุ่งวัดลักษณะเดียวกัน และหลักฐานการเปรียบเทียบความสัมพันธ์กับผลการวัดด้วยเครื่องมืออื่นที่มุ่งวัดลักษณะเดียวกัน และลักษณะที่ต่างกัน เป็นต้น มีวิธีการสำหรับหาหลักฐานเชิงประจักษ์เพื่อใช้สนับสนุนความตรงเชิงทฤษฎีหลายวิธี วิธีที่นิยมใช้กันมีดังนี้

1) วิธีเปรียบเทียบคะแนนระหว่างกลุ่มที่ทราบผล (Comparing the scores of known groups)

ในกรณีที่มีความเชื่อมั่นตามทฤษฎีว่า คะแนนผลการวัดลักษณะที่สนใจนั้นจะมีความแตกต่างกันระหว่างกลุ่มผู้สอบที่ทราบแน่ชัดว่า มีลักษณะสำคัญบางประการแตกต่างกัน เช่น แตกต่างกันตามกลุ่มอายุ กลุ่มความสามารถกลุ่มที่ผ่านประสบการณ์เฉพาะอย่าง กลุ่มที่มีการปรับตัวได้ดี / ปรับตัวไม่ดี กลุ่มที่ประสบความสำเร็จ / ไม่ประสบความสำเร็จ เป็นต้น จึงน่าจะหาหลักฐานส่วนหนึ่งใช้สนับสนุน ความตรงเชิงทฤษฎีได้ด้วย การเปรียบเทียบคะแนนที่วัดได้ระหว่างกลุ่มที่ทราบแน่ชัดแล้วว่ามีลักษณะที่มุ่งวัดนั้นแตกต่างกัน (known groups) ถ้าเครื่องมือสามารถวัดลักษณะที่น่าสนใจนั้นได้ ผลการวัดจะต้องมีความแตกต่างระหว่างกลุ่ม

จากการศึกษาทฤษฎีเพื่อสร้างแบบสอบวัดลักษณะต่าง ๆ บางทฤษฎีอาจให้ทิศทางความแตกต่างระหว่างคะแนนของกลุ่มผู้สอบ เช่น สมมติว่าทฤษฎีได้คาดหมายไว้ว่า คะแนนผลสัมฤทธิ์เกี่ยวกับ X จะแตกต่างกันระหว่างกลุ่มผู้เรียนที่ใช้หลักสูตร A และกลุ่มผู้เรียนที่ใช้หลักสูตร B คะแนนความสามารถด้าน Y จะแตกต่างกันระหว่างกลุ่มอายุของผู้สอบ คะแนนความสามารถในการปรับตัวจากแบบสอบ Z ที่สร้างขึ้นตามทฤษฎี ย่อมจะใช้จำแนกกลุ่มบุคคลที่ปรับตัวได้ดีกับกลุ่มบุคคลที่มีปัญหาในการปรับตัวออกจากกันได้ ในกรณีของแบบสอบ Z ถ้าแบบสอบ Z วัดความสามารถในการปรับตัวได้ เมื่อนำไปใช้ทดสอบกับกลุ่มนักเรียนที่ทราบแน่ชัดว่าสามารถปรับตัวได้ดีกับกลุ่มที่มีปัญหาในการปรับตัว กลุ่มทั้งสองควรจะต้องได้คะแนนจากแบบสอบ Z แตกต่างกันตามทิศทางที่ควรจะเป็นตามทฤษฎี เป็นต้น

2) วิธีการเปรียบเทียบคะแนนจากการทดลอง (Comparing scores from experiment)

คะแนนจากเครื่องมือวัดลักษณะใดก็ตาม คาดว่าน่าจะเปลี่ยนแปลงได้ตามเงื่อนไขของการจัดกระทำตามการทดลอง อาจมีการเปลี่ยนแปลงระหว่างกลุ่มทดลอง ก่อนหลังได้รับการจัดกระทำตามตัวแปรทดลอง เช่น ตามทฤษฎีคาดหมายว่า คะแนนความวิตกกังวลของบุคคลน่าจะเปลี่ยนแปลงไปตามสถานการณ์ ถ้ามีการสร้างสถานการณ์ให้เกิดความวิตกกังวลระดับต่าง ๆ ขึ้นมา เมื่อใช้แบบสอบวัดความวิตกกังวล ก็น่าจะวัดได้คะแนนความวิตกกังวลที่ระดับต่าง ๆ สอดคล้องกับสถานการณ์ตามคำทำนายของทฤษฎี เป็นต้น

จากทฤษฎีที่เกี่ยวกับลักษณะที่มุ่งวัด จึงสามารถทำนายถึงทิศทางและปริมาณการเปลี่ยนแปลงของลักษณะที่มุ่งวัดภายใต้เงื่อนไขต่าง ๆ ของการทดลองได้ ถ้าผลที่ได้จากการทดลองสอดคล้องหรือยืนยันคำทำนายของทฤษฎี ผลที่ได้จะเป็นหลักฐานส่วนหนึ่งสำหรับใช้สนับสนุนความตรงเชิงทฤษฎีของแบบสอบได้

3) วิธีวิเคราะห์เมตริกซ์พหุลักษณะ– พหุวิธี (Multitrait – Multimethod :MTMM)

แคมพ์เบล และ ฟิสค์ (Campbell and Fiske, 1959 อ้างถึงใน ศิริชัย กาญจนวาสี, 2544) ได้เสนอวิธีตรวจสอบความตรงเชิงทฤษฎีด้วยการวิเคราะห์เมตริกซ์พหุลักษณะ – พหุวิธี ซึ่งเป็นวิเคราะห์ความสัมพันธ์ระหว่างการวัดลักษณะหลายลักษณะ (Multitrait) โดยใช้วิธีการวัดหลายวิธี (Multimethod)

ขั้นตอนการใช้วิธีวิเคราะห์เมตริกซ์พหุลักษณะพหุวิธี

วิธีพหุลักษณะ – พหุวิธี (MTMM) นี้ มุ่งการตรวจสอบความเหมาะสมของเครื่องมือหลายประเภทสำหรับวัดลักษณะหนึ่งที่สนใจ มากกว่าที่จะเป็นการตรวจสอบยืนยันความสัมพันธ์ระหว่างการวัดลักษณะหลายอย่างตามคำทำนายของทฤษฎี การใช้วิธีนี้สำหรับตรวจสอบความตรงเชิงทฤษฎี มีขั้นตอนการดำเนินงานดังนี้

(1) คัดเลือกวิธีการวัด ที่เหมาะสมอย่างน้อย 2 วิธีสำหรับวัด ลักษณะที่สนใจ

(2) คัดเลือก ลักษณะอื่นที่มีความแตกต่างจากลักษณะที่สนใจ และสามารถวัดได้โดยวิธีการวัดตามข้อ (1)

(3) สร้างเครื่องมือ วัดลักษณะทั้งหลายด้วยวิธีการหลายอย่างตามที่กำหนดไว้ใน (1) และ (2) เช่น ถ้ามี p ลักษณะ และ q วิธี จะมีเครื่องมือทั้งหมด pq ชุด เป็นต้น

(4) เก็บรวบรวมข้อมูล นำเครื่องมือแต่ละชนิดไปวัดลักษณะแต่ละลักษณะของกลุ่มตัวอย่าง

(5) คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ ระหว่างคะแนนการวัดแต่ละลักษณะ โดยใช้เครื่องมือแต่ละชนิด ทำให้ได้เมตริกซ์พหุลักษณะ – พหุวิธีขนาด pxq

(6) แปลความหมาย ของค่าสัมประสิทธิ์สหสัมพันธ์ในเมตริกซ์พหุลักษณะพหุวิธี วิธีนี้สามารถนำมาใช้ได้ เมื่อมีการวัดลักษณะอย่างน้อย 2 ลักษณะโดยใช้วิธีการวัดอย่างน้อย 2 วิธี ตัวอย่าง เช่น การวัดลักษณะที่แตกต่างกัน 3 ลักษณะ ได้แก่ A, b และ C โดยใช้วิธีการทดสอบที่แตกต่างกัน 3 วิธี ได้แก่ วิธีที่ 1, 2 และ 3 เมื่อนำเครื่องมือหรือแบบสอบทั้ง 9 ฉบับไปทดสอบกับกลุ่มตัวอย่างเดียวกัน ค่าคะแนนที่ได้มาคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่ได้จากเครื่องมือทั้ง 9 ฉบับ จะได้เมตริกซ์สหสัมพันธ์ ซึ่งเรียกว่า เมตริกซ์พหุลักษณะ – พหุวิธี ในกรณีนี้เมตริกซ์จะมีขนาด 9 x 9 ดังตารางที่ 2.2

ตารางที่ 2.2 ส่วนประกอบของเมตริกซ์พหุลักษณะ – พหุวิธีขนาด 3 x 3 (ศิริชัย กาญจนวาสี, 2544)

วิธีวัด	คุณลักษณะ	A				B				C		
		1	2	3		1	2	3		1	2	3
A	1	[ ]										
	2	…..	[ ]									
	3	…..	…..	[ ]								
B	1	( )	…	….		[ ]						
	2	…..	( )	…..		…..	[ ]					
	3	…..	…..	( )		…..	…..	[ ]				
C	1	( )	…	….		( )	…	….		[ ]		
	2	…..	( )	…..		…..	( )	…..		…..	[ ]	
	3	…..	…..	( )		…..	…..	( )		…..	…..	[ ]

ส่วนประกอบของเมตริกซ์พหุลักษณะ – พหุวิธี (ศิริชัย กาญจนวาสี, 2544)

ค่าสัมประสิทธิ์สหสัมพันธ์ในเมตริกซ์พหุลักษณะ–พหุวิธี สามารถจำแนกได้เป็น 3 ประเภทดังนี้

(1) สัมประสิทธิ์ความเที่ยง (Reliability coefficient)

สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากการวัดลักษณะเดียวกัน โดยใช้วิธีการเดียวกัน (Monotrait – Monomethod ; MTMM) เปรียบเหมือนการวัดลักษณะเดียวกันโดยการวัดซ้ำด้วยวิธีการเดิม เมื่อนำคะแนนจากการวัดลักษณะเดียวกันซ้ำมาหาความสัมพันธ์ ผลที่ได้ก็คือสัมประสิทธิ์ความเที่ยงนั่นเอง จากตาราง สัมประสิทธิ์ความเที่ยงเป็นค่าที่อยู่ในแนวทแยงมุมหลักกึ่งกลางเมตริกซ์ ซึ่งแสดงด้วยสัญลักษณ์ [ ]

(2) สัมประสิทธิ์ความตรง (Validity coefficient)

สัมประสิทธิ์สัมพันธ์ระหว่างคะแนนจากการวัดลักษณะเดียวกัน โดยใช้วิธีต่างกัน (Monotrait – Heteromethod ; MTHM) ซึ่งเป็นค่าสัมประสิทธิ์ความตรงนั่นเอง จากตาราง สัมประสิทธิ์ความตรงเป็นค่าที่อยู่ในสามเหลี่ยมกรอบทึบ

(3) สัมประสิทธิ์สหสัมพันธ์ (Correlation coefificient)

สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากการวัดลักษณะต่างกันซึ่งมีอยู่ในเมตริกซ์ พหุลักษณะ – พหุวิธี สำหรับเปรียบเทียบการจำแนกกับสัมประสิทธิ์ความเที่ยงและสัมประสิทธิ์ความตรง สัมประสิทธิ์สหสัมพันธ์เชิงจำแนกมี 2 ประเภทดังนี้

(3.1) สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากการวัด ลักษณะต่างกัน โดยใช้วิธีเดียวกัน (Heterotrait – Monomethod ; HTMM) จากตาราง สัมประสิทธิ์สหสัมพันธ์ประเภทนี้เป็นค่าแนวทะแยงที่แสดงด้วยสัญลักษณ์

(3.2) สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากการวัด ลักษณะต่างกัน โดยใช้ วิธีต่างกัน (Heterotrtait – Heteromethod ; HTHM) จากตาราง สัมประสิทธิ์สหสัมพันธ์ ประเภทนี้เป็นค่าที่อยู่ในรูปสามเหลี่ยมกรอบประ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การแปลผลเมตริกซ์พหุลักษณะ – พหุวิธี (ศิริชัย กาญจนวาสี, 2544)

แคมพ์เบล และฟิสด์ ได้ให้แนวทางของการแปลผลเมตริกซ์พหุลักษณะ – พหุวิธี ได้ว่า วิธีการวัดหลายวิธีจะลู่เข้า (converge) สำหรับการวัดลักษณะเดียวกันกล่าวคือ คะแนนจากการวัดลักษณะเดียวกัน ถึงแม้จะใช้วิธีการวัดต่างกันย่อมมีความสัมพันธ์กันสูง และวิธีการวัดจะต้องจำแนก (discriminate) สำหรับการวัดลักษณะต่างกัน กล่าวคือ คะแนนจากการวัดคุณลักษณะต่างกันไม่ว่าจะใช้วิธีการวัดเดียวกันหรือต่างกัน ย่อมมีความสัมพันธ์กันต่ำ แต่ถ้าความสัมพันธ์ไม่เป็นไปตามนี้ ทำให้สงสัยได้ว่ามีปัญหาเกิดขึ้นกับวิธีการวัด (เครื่องมือ) หรือลักษณะที่มุ่งวัดการแปลผลเมตริกซ์พหุลักษณะ – พหุวิธีสามารถกระทำได้ดังนี้

(1) สัมประสิทธิ์ความเที่ยงซึ่งเป็นค่าที่อยู่ในแนวทแยงหลักกึ่งกลางเมตริกซ์ จะต้องมีค่าสูง มีนัยสำคัญทางสถิติและมีขนาดใหญ่ที่จะนำไปใช้พิจารณาความตรงเชิงทฤษฎีต่อไป

(2) สัมประสิทธิ์ความตรง ซึ่งเป็นค่าที่อยู่ในสามเหลี่ยมกรอบทึบ ควรมีค่าสูง มีนัยสำคัญและมีขนาดใหญ่พอสำหรับใช้เป็นหลักฐานสนับสนุนความตรงแบบลู่เข้า(convergent validity) ค่าสัมประสิทธิ์ความตรงนี้อาจมีค่าอ่อนตัวลงเนื่องจากความคลาดเคลื่อนจากการวัด หรือปัญหาความเที่ยงของเครื่องมือ

(3) สัมประสิทธิ์สัมพันธ์ระหว่างคะแนนจากการวัดลักษณะต่างกัน ซึ่งเป็นค่าที่อยู่ในแนวทแยง ( ) และสามเหลี่ยมกรอบประ ควรมีค่าต่ำ โดยต่ำกว่าค่าสัมประสิทธิ์ความเที่ยงและสัมประสิทธิ์ความตรง ซึ่งเป็นหลักฐานแสดงถึงความตรงเชิงจำแนก (discriminant validity)

(4) แบบแผนของสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนน จากการวัดลักษณะเดียวกันหรือการวัดลักษณะต่างกัน ควรเหมือนกัน

วิธีวิเคราะห์องค์ประกอบ (Factor Analysis)

หลักฐานอย่างหนึ่งที่สามารถนำมาใช้สนับสนุนความตรงเชิงทฤษฎีได้แก่ ความตรงตามองค์ประกอบ (Factorial validity) ซึ่งเป็นคุณสมบัติของการวัดได้ตรงตามตัวประกอบที่มุ่งวัด (Nunally, 1978 ; Allen and Yen , 1979 อ้างถึงใน ศิริชัย กาญจนวาสี, 2544)

ความตรงตามองค์ประกอบ สามารถศึกษาได้ด้วย การวิเคราะห์องค์ประกอบ (factor analysis) ซึ่งเป็นเทคนิคทางสถิติสำหรับวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรที่สังเกตค่าได้ เพื่อหาลักษณะร่วมกันของชุดตัวแปรเหล่านั้น ลักษณะร่วมกันนี้เรียกว่า องค์ประกอบ (factor) องค์ประกอบเป็นลักษณะที่คาดว่ามี อิทธิพลต่อคะแนนที่ได้จากกลุ่มตัวแปร หรือเป็นลักษณะที่ใช้อธิบายความผันแปรร่วมของกลุ่มตัวแปร อย่างไรก็ตามผลการวิเคราะห์องค์ประกอบจำเป็นต้องแปลผลในบริบทของทฤษฎีที่เกี่ยวข้องกับลักษณะที่ทำการวิเคราะห์นั้น เนื่องจากการ

กำหนดโครงสร้างของคุณลักษณะร่วมนี้ จำเป็นต้องได้รับการแนะนำโดยทฤษฎีที่สำคัญอันเป็นที่ยอมรับ มิเช่นนั้นอาจนำไปสู่ข้อสรุปที่ผิดพลาดเกี่ยวกับความตรงตามลักษณะที่มุ่งวัดได้

องค์ประกอบเป็นตัวแปรเชิงสมมติฐานที่ไม่สามารถสังเกตได้โดยตรง แต่คาดว่าเป็นลักษณะหรือโครงสร้างที่เกิดจากการเกาะกลุ่มของตัวแปรที่มีความสัมพันธ์สูง จากการวิเคราะห์องค์ประกอบทำให้ได้ องค์ประกอบ ซึ่งสามารถใช้เป็นหลักฐานตรวจสอบความตรงเชิงทฤษฎีของเครื่องมือหรือแบบสอบได้ใน 2 ลักษณะได้แก่ (1) แบบสอบนั้นมุ่งวัดลักษณะได้สอดคล้องกับโครงสร้างทางทฤษฎีของลักษณะที่มุ่งวัดนั้นเพียงไร (วัดส่วนประกอบได้ครอบคลุมโครงสร้างทางทฤษฎีของลักษณะที่สนใจ) และ (2) แบบสอบนั้นมุ่งวัดลักษณะได้ตรงตามลักษณะที่ต้องการวัดนั้นเพียงไร (วัดองค์ประกอบร่วมได้ตรงตามลักษณะที่สนใจ)

โมเดลสำหรับวิเคราะห์ตัวประกอบที่ใช้อยู่ในปัจจุบันมีอยู่ 2 โมเดล ได้แก่ โมเดลการวิเคราะห์องค์ประกอบเชิงสำรวจ (Explorator Factor Model) และโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Model)

จากที่กล่าวมาข้างต้น สามารถสรุปประเภทของความตรง ความหมายและวิธีการตรวจสอบ ได้ดังนี้

ตารางที่ 2.3 ประเภทของความตรง ความหมาย และวิธีการตรวจสอบ (ศิริชัย กาญจนวาสี, 2544)

ประเภท	ความหมาย	วิธีการตรวจสอบ
1. ความตรงตามเนื้อเรื่อง (content validity)	ความสามารถในการวัดกลุ่มตัวอย่างเนื้อเรื่อง วัดได้ครอบคลุม และเป็นตัวแทนของมวลเนื้อเรื่องหรือประสบการณ์ที่มุ่งวัด	1. ให้ผู้เชี่ยวชาญตรวจความเหมาะสมของนิยามและขอบเขตของมวลเนื้อเรื่องหรือประสบการณ์ที่มุ่งวัด 2. ตรวจสอบกลุ่มตัวอย่างเนื้อเรื่องหรือพฤติกรรมที่นำมาใช้วัดในเครื่องมือว่ามีความครอบคลุมเนื้อเรื่องหรือประสบการณ์ทั้งหมดหรือไม่ เพียงไร 3. เปรียบเทียบสัดส่วนของข้อคำถามว่ามีความสอดคล้องกับน้ำหนักความสำคัญของแต่ละลักษณะเนื้อเรื่องที่มุ่งวัดมากน้อยเพียงไร
2. ความตรงตามเกณฑ์สัมพันธ์ (criterion related validity)	ความสามารถในการวัดลักษณะที่สนใจได้สอดคล้องกับเกณฑ์ภายนอก	คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากเครื่องมือกับคะแนนจากเกณฑ์ภายนอก ซึ่งวัดได้

ประเภท	ความหมาย	วิธีการตรวจสอบ
		จากเครื่องมืออิสระอื่นที่เชื่อถือได้
2.1 ความตรง ตามสภาพหรือร่วมสมัย (concurrent validity)	ความสามารถในการวัดลักษณะที่สนใจได้ตรงตามสมรรถนะของสิ่งนั้นใน สภาพปัจจุบัน	คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากเครื่องมือกับคะแนนที่วัดได้จากเครื่องมือมาตรฐานอื่น ซึ่งสามารถวัดสิ่งนั้นได้ในสภาพปัจจุบัน
2.2 ความตรงเชิงทำนาย (predictive validity)	ความสามารถในการวัดลักษณะที่สนใจได้ตรงตามสมรรถนะของสิ่งนั้นที่จะเกิดขึ้นในอนาคต	คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่วัดได้จากเครื่องมือกับคะแนนที่วัดได้จากเครื่องมือมาตรฐานอื่นซึ่งสามารถวัดสิ่งนั้นได้ในเวลาต่อมาหรือในอนาคต
3. ความตรงเชิงทฤษฎีหรือโครงสร้าง (construct validity)	ความสามารถในการวัดได้ตรงตามลักษณะที่มุ่งวัด โดยผลการวัดมีความสอดคล้องกับโครงสร้างและความหมายทางทฤษฎีของลักษณะที่มุ่งวัดนั้น	ศึกษาความสัมพันธ์ระหว่างผลการวัดที่ได้จากเครื่องมือกับโครงสร้างและคำทำนายทางทฤษฎีของลักษณะที่มุ่งวัดโดยอาศัยข้อสนับสนุนเชิงสะสมของหลักฐานจากวิธีการวิเคราะห์ต่างๆ เช่น 1) วิธีเปรียบเทียบคะแนนระหว่างกลุ่มที่ทราบผล 2) วิธีเปรียบเทียบคะแนนจากการทดลอง 3) วิธีวิเคราะห์เมตริกซ์พหุลักษณะ พหุวิธี 4) วิธีวิเคราะห์องค์ประกอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4 ทฤษฎีการสรุปอ้างอิง(Generalizability Theory)

ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Theory) หรือเรียกว่า G -Theory เป็นทฤษฎีทางสถิติของการวิเคราะห์ความน่าเชื่อถือของผลการวัดในสถานการณ์ของการวัดลักษณะต่าง ๆ ที่เป็นเป้าหมายของการนำเครื่องมือไปใช้ ซึ่ง G-Theory เป็นทั้งวิธีประเมินความน่าเชื่อถือหรือความเที่ยงของเครื่องมือวัดผล และกลยุทธ์ของการออกแบบการวัดให้ได้ผลของการวัดที่มีความน่าเชื่อถือหรือมีความเที่ยงสูงถึงระดับที่ต้องการ เพื่อนำผลไปใช้เป็นสารสนเทศสำหรับการตัดสินใจได้อย่างมีประสิทธิภาพ

แนวคิดพื้นฐานของ G – theory (ศิริชัย กาญจนวาสี, 2544)

ครอนบาค และคณะ (Cronbach & other , 1972 อ้างถึงใน ศิริชัย กาญจนวาสี, 2544) ได้เสนอทฤษฎีสำหรับวิเคราะห์ความน่าเชื่อถือของผลการวัดพฤติกรรม (The Dependability of behavioral Measurements) สำหรับสถานการณ์ของการวัดผลลักษณะต่างๆ ซึ่งต่อมาเป็นที่รู้จักกันชื่อของทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด ทฤษฎีนี้ได้ขยายแนวความคิดของความเที่ยง ตามทฤษฎีการทดสอบแบบดั้งเดิม (CTT) ให้สามารถสรุปผลความเที่ยงในสถานการณ์หรือเงื่อนไขการทดสอบลักษณะต่างๆ ได้

ทฤษฎีการทดสอบแบบดั้งเดิมมิได้ให้ความสนใจต่อสถานการณ์หรือเงื่อนไขของการวัดและปฏิสัมพันธ์ระหว่าง ผู้สอบ กับเงื่อนไขของการวัดที่สามารถส่งผลต่อความน่าเชื่อถือของผลการวัด แต่ G - Theory ได้ในแนวคิดที่แยกส่วนความคลาดเคลื่อน (error) จากหลายแหล่ง (multiple error sources) ประกอบด้วยแหล่งความคลาดเคลื่อนอย่างเป็นระบบ (systematic source) และความคลาดเคลื่อนสุ่ม (random source) มิใช่เป็นความคลาดเคลื่อนรวมเพียงแหล่งเดียวเหมือนอย่างทฤษฎีการทดสอบแบบดั้งเดิม

ความรู้พื้นฐานเกี่ยวกับโครงสร้างของการศึกษาทฤษฎีการสรุปอ้างอิง ได้เสนอคำศัพท์ที่ใช้เฉพาะในทฤษฎีการสรุปอ้างอิง ไว้ดังนี้

1. ประชากร (population) หมายถึงสิ่งที่เป็นเป้าหมายของการวัดหรือสิ่งที่มุ่งวัดทั้งหมด (object of measurement) ในสถานการณ์ของการทดสอบโดยทั่วไป สิ่งที่มุ่งวัดมักได้แก่ ผู้สอบ บุคคล นักเรียน หรือนักเรียนทั้งชั้น

2. องค์ประกอบ (facet) หมายถึงชุดของเงื่อนไขของการวัดที่มีลักษณะคล้ายกัน หรือกลุ่มเงื่อนไขของการวัด ซึ่งเป็นองค์ประกอบที่คาดว่ามีผลต่อการวัดความคลาดเคลื่อน เช่น

องค์ประกอบของข้อสอบ (item facet) องค์ประกอบของจำนวนครั้งของการสอบ (occation facet) องค์ประกอบของผู้ตรวจข้อสอบ (rater facet)

3. เงื่อนไขของการวัด (condition of measurement) หมายถึงระดับขององค์ประกอบที่ทำให้ได้ค่าสังเกตแต่ละค่าขึ้นมาในการวัดครั้งหนึ่งๆ เช่น ฟาเซทของจำนวนผู้ตรวจ อาจกำหนดจำนวนระดับเป็น 1 , 2 และ 3 คน ฟาเซทความยาวของแบบสอบ อาจกำหนดความยาวของแบบสอบ เป็น 10 , 20 และ 30 ข้อ เป็นต้น

4. รูปแบบความสัมพันธ์ระหว่างองค์ประกอบ หมายถึงองค์ประกอบที่เป็นสิ่งที่ชี้ถึงแหล่งความแปรปรวนของรูปแบบการวัด ในแบบจำลองการวิเคราะห์ความแปรปรวน ซึ่งรูปแบบความสัมพันธ์นี้แบ่งออกเป็น 3 ลักษณะ คือ

4.1 ความสัมพันธ์แบบไขว้ (Crossed) หมายถึงความสัมพันธ์แต่ละลักษณะที่มีแต่ละระดับของสิ่งที่ถูกวัด ถูกวัดภายใต้เงื่อนไขเดียวกันทั้งหมด สัญลักษณะคือ “×” อ่านว่า “crossed with” เช่น กำหนดรูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็น p× i×r หมายถึง นักเรียน (p) ทำข้อสอบ (i) ทุกข้อ และผู้ตรวจ (r) ตรวจข้อสอบของนักเรียนทุกคน ทุกๆ ข้อ

4.2 ความสัมพันธ์แบบแฝง (Nested) หมายถึงความสัมพันธ์ในลักษณะที่แต่ละระดับของสิ่งที่ถูกวัด ถูกวัดภายใต้เงื่อนไขเดียวกันทั้งหมด สัญลักษณะคือ “:” อ่านว่า “nested within” เช่น รูปแบบความสัมพันธ์องค์ประกอบเป็น i:r หมายถึงผู้ตรวจข้อสอบ (r) ตรวจข้อสอบ (i) ต่างข้อกัน

4.3 ความสัมพันธ์แบบผสม (Confounded) หมายถึงความสัมพันธ์ที่มีทั้งความสัมพันธ์แบบไขว้และความสัมพันธ์แบบแฝงปนกันอยู่ เช่น p x (i :t) หมายถึง นักเรียน (p) ทำข้อสอบ (i) บางข้อที่อยู่ในแบบทดสอบ (t)

5. เอกภพ (Universe) หมายถึงเงื่อนไขของการวัด (condition of measurement) ทั้งหมด กลุ่มเงื่อนไขของการวัดหรือเรียกว่า ฟาเซท ซึ่งเป็นองค์ประกอบที่คาดว่ามีผลต่อความคลาดเคลื่อนของการวัด เช่น ความยาวของแบบสอบ รูปแบบของข้อสอบ จำนวนครั้งของการทดสอบ จำนวนผู้ตรวจทั้งหมด เป็นต้น

คะแนนเอกภพ (Universe Score) หมายถึงคะแนนของเงื่อนไขการวัดทั้งหมดแต่ละองค์ประกอบซึ่งคล้ายกับคะแนนจริง (Ideal Score) ของสิ่งที่วัดในทฤษฎีการวัดแบบดั้งเดิม

6. เอกภพของการสังเกตที่ยอมรับได้ (Universe of Admissible Observation) หมายถึงกลุ่มเงื่อนไขของการวัดที่เป็นไปได้ ซึ่งสามารถวัดหรือสังเกตได้ในแต่ละฟาเซท เช่น ประกอบด้วยฟาเซทข้อสอบและฟาเซทของผู้ตรวจให้คะแนน

7. เอกภพของการสรุปอ้างอิง (Universe of Generalization) หมายถึงสถานการณ์หรือเงื่อนไขการวัดอันเป็นเป้าหมายของการสรุปอ้างอิง กล่าวได้ว่า เป็นการวัดที่ครอบคลุมเงื่อนไขที่สนใจทั้งหมด หรือเงื่อนไขในเอกภพของการสังเกตที่ยอมรับได้ทั้งหมด ซึ่งอาจประกอบไปด้วยเซ็ตย่อยของเงื่อนไขในเอกภพของการสังเกตได้

8. การศึกษาเพื่อการสรุปอ้างอิง (Generalizability Study: G Study) หมายถึงการวิเคราะห์การสรุปอ้างอิง เพื่อประมาณค่าความแปรปรวนจากแหล่งต่างๆ ของรูปแบบการวัด โดยเริ่มจากการกำหนดองค์ประกอบต่างๆ กำหนดความสัมพันธ์ขององค์ประกอบ และกำหนดเอกภพของการสังเกตที่ยอมรับได้ และมีการวิเคราะห์ความแปรปรวน การประมาณค่าความแปรปรวนจากแหล่งต่างๆ

9. การศึกษาเพื่อการตัดสินใจ (Decision Study) หมายถึงการศึกษาเพื่อประเมินวิธีการวัดที่เหมาะสม เพื่อการตัดสินใจหรือลงข้อสรุป เป็นขั้นสุดท้ายของการวิเคราะห์การสรุปอ้างอิง โดยเริ่มจากผู้วิจัยต้องระบุสิ่งที่ต้องการวัด เอกภพของการอ้างอิง จำนวนเงื่อนไขของการวัด หรือขนาดของกลุ่มตัวอย่างของแต่ละองค์ประกอบ และรูปแบบของการวัด แล้วหาค่าความแปรปรวนจากการศึกษา ประมาณค่าความแปรปรวน หาค่าสัมประสิทธิ์หรือความเที่ยงของแบบทดสอบ และสุดท้ายปรับรูปแบบการวัดเพื่อหาวิธีที่เหมาะสมเพื่อให้เกิดประโยชน์สูงสุดต่อไป

ขนาดของกลุ่มตัวอย่าง ในการศึกษาเพื่อการตัดสินใจ ไม่จำเป็นต้องมีขนาดเหมือนกับขนาดของกลุ่มตัวอย่าง ที่ใช้ในการศึกษาเพื่อการสรุปอ้างอิง แต่ขึ้นอยู่กับการนำไปใช้เพื่อการตัดสินใจเกี่ยวกับสิ่งที่จะวัดจริงๆ ซึ่งจะถูกกำหนดขึ้นโดยมีขนาดของกลุ่มตัวอย่างใน การศึกษาเพื่อการสรุปอ้างอิงเป็นพื้นฐาน

10. ความแปรปรวนของคะแนนความคลาดเคลื่อนสัมบูรณ์ (Absolute Error Variance: σ^2_{Abs} หรือ σ^2_{Δ}) หมายถึงความแปรปรวนของผลต่างระหว่างคะแนนสังเกตได้กับคะแนนเอกภพ ($\mu_p - x_p$) ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่างๆ ยกเว้นความแปรปรวนของแหล่งบุคคล ($\sigma^2_{\mu_p}$ หรือ σ^2_{p})

11. ความแปรปรวนของความคลาดเคลื่อนแบบสัมพัทธ์ (Relative Error Variance: σ^2_{Rel} หรือ σ^2_{δ}) หมายถึงความแปรปรวนของผลต่างระหว่างคะแนนสังเกตได้ และคะแนนเอกภพ ($\mu_p - X_p$) ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่างๆ ที่มีปฏิสัมพันธ์ร่วมกับผู้สอบ (p)

12. สัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient : p^2) หมายถึงอัตราส่วนของความแปรปรวนของคะแนนเอกภพและความแปรปรวนของคะแนนสังเกตที่คาดหวังซึ่งค่า

สัมประสิทธิ์การสรุปอ้างอิงที่ประมาณค่าได้จาก กำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนเอกภพและคะแนนสังเกตได้

12.1 สัมประสิทธิ์การสรุปอ้างอิง สำหรับการวัดที่มีการตัดสินใจในเชิงสัมพัทธ์หรือการวัดแบบ อิงกลุ่ม เป็นความคลาดเคลื่อนแบบสัมพัทธ์ ค่าสัมประสิทธิ์การสรุปอ้างอิงประมาณค่าได้จาก อัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและผลบวกของความแปรปรวนของคะแนนเอกภพ กับความแปรปรวนความคลาดเคลื่อนสัมพัทธ์

12.2 สัมประสิทธิ์การสรุปอ้างอิง สำหรับการวัดที่มีการตัดสินใจแบบสัมบูรณ์หรือการวัดแบบ อิงเกณฑ์ เป็นความคลาดเคลื่อนในการวัดแบบสัมบูรณ์ ค่าสัมประสิทธิ์การสรุปอ้างอิงประมาณค่าได้จากอัตราส่วนความแปรปรวนของคะแนนเอกภพและผลบวกของความแปรปรวนของคะแนนเอกภพกับความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์

การวิเคราะห์สัมประสิทธิ์การสรุปอ้างอิง

การวิเคราะห์สัมประสิทธิ์การสรุปอ้างอิง (Cronbach & other , 1972 ; Brennan, 1983 อ้างถึงใน ศิริชัย กาญจนวาสี, 2544) ได้เสนอขั้นตอน การวิเคราะห์สัมประสิทธิ์การสรุปอ้างอิงออกเป็น 2 ขั้นตอน คือขั้นตอนแรกเป็นการศึกษาเพื่อการสรุปอ้างอิง หรือการศึกษา G (Generalizability Study: G-Study) และขั้นตอนที่สองเป็นการศึกษาเพื่อการตัดสินใจ หรือการศึกษา D (Decision Study: D-Study) สำหรับหลักการดำเนินงานในแต่ละขั้นตอน มีดังต่อไปนี้

ขั้นที่ 1 การศึกษาเพื่อการสรุปอ้างอิง

การศึกษาเพื่อการสรุปอ้างอิง เป็นกระบวนการที่มุ่งประมาณความแปรปรวนจากแหล่งต่างๆ ภายใต้เอกภพของคะแนนสังเกตที่ยอมรับได้ โดยมีขั้นตอนดังนี้

1. กำหนดสิ่งที่จะวัด เช่น กำหนดสิ่งที่ต้องการวัด คือ ความสามารถด้านการเรียนของ นักเรียน

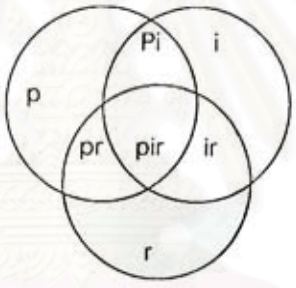
2. กำหนดฟาเซทของการวัด เช่น กำหนดให้มี 2 ฟาเซท คือ ข้อสอบ (i) และผู้ตรวจให้คะแนน (r)

3. กำหนดเอกภพของการสังเกตที่ยอมรับได้ ในที่นี้ประกอบด้วยฟาเซทข้อสอบฟาเซทของผู้ตรวจให้คะแนน

4. กำหนดความสัมพันธ์ระหว่างฟาเซท ว่าจะเป็นแบบ crossed ซึ่งแทนด้วย “x” หรือ nested แทนด้วย “.” ถ้าต้องการศึกษาให้ผู้ตรวจทุกคนข้อสอบทุกข้อก็เป็นแบบ crossed คือ ixr แต่ถ้าให้ผู้ตรวจแต่ละคนตรวจชุดข้อสอบต่างกันก็ถือว่าเป็นแบบ nested คือ i:r ทั้งนี้แล้วแต่แล้วแต่ความสนใจของผู้ศึกษา ซึ่งโดยหลักการแล้วถ้าเป็นแบบ crossed design จะทำให้ได้

จำนวนแหล่งความแปรปรวนมากกว่าแบบ nested design ซึ่งมีประโยชน์ที่จะทำให้ทราบว่า ความแปรปรวนแหล่งใดมีผลต่อสัมประสิทธิ์การสรุปอ้างอิงที่มีค่ามากกว่า nested design ในกรณี ของ 1-facet design แต่ในแง่ปฏิบัตินั้น nested design จะสามารถใช้ได้ง่ายกว่า กล่าวคือ การให้ผู้ตรวจแต่ละคนตรวจข้อสอบต่างข้อกัน โดยสรุปแล้วในขั้น G-Study นั้น ควรใช้ crossed design เพราะขั้นนี้มุ่งประมาณค่าความแปรปรวนจากแหล่งต่างๆ อย่างเต็มที่ สำหรับในขั้น D-Study การเลือกใช้รูปแบบใดขึ้นอยู่กับความสนใจที่ต้องการศึกษา

5. การเก็บรวบรวมข้อมูล โดยการสุ่มตัวอย่างมาศึกษา เช่น ในที่นี้สุ่มตัวอย่างนักเรียน n_p คน ตัวอย่างข้อสอบ n_i และตัวอย่างผู้ตรวจ n_r คน จะเป็นการศึกษาแบบ G-Study pxix r design ในกรณีนี้จะประกอบด้วย แหล่งความแปรปรวน 7 แหล่ง เป็นผลหลัก (main effect) 3 แหล่ง คือ ผลจากนักเรียน ผลจากข้อสอบและผลจากผู้ตรวจข้อสอบและเป็นผลของปฏิสัมพันธ์อีก 4 แหล่ง ได้แก่ ปฏิสัมพันธ์ระหว่างนักเรียนกับข้อสอบ นักเรียนกับผู้ตรวจข้อสอบ ข้อสอบกับข้อสอบ และนักเรียนกับข้อสอบและผู้ตรวจ



แผนภาพที่ 2.6 แผนภูมิแสดงองค์ประกอบความแปรปรวนของรูปแบบ pxix r เมื่อตัวประกอบที่ศึกษาทั้งหมดเป็นตัวประกอบสุ่ม

จากภาพที่ 6 จะเห็นว่า มีผลหลัก (Main Effects) อยู่ 3 ค่า คือ 1) ผลของผู้สอบ (P) 2) ผลของข้อสอบ (i) และ 3) ผลของผู้ตรวจ (r) ส่วนผลร่วมหรือปฏิสัมพันธ์สองระดับมี 3 ค่า คือ 1) ผลร่วมของผู้สอบและข้อสอบ (pi) 2) ผลร่วมของผู้สอบกับผู้ตรวจ (pr) 3) ผลร่วมของข้อสอบและ ผู้ตรวจ (ir) และผลร่วมทั้งหมดของผู้สอบ ข้อสอบและผู้ตรวจ (pir)

ค่าความแปรปรวนของแหล่งต่างๆ เหล่านี้ คำนวณมาจากค่าเฉลี่ยกำลังสอง (Mean Square) โดยใช้การวิเคราะห์ความแปรปรวนแบบ Factorial Design p x i x r ในการประมาณค่าความแปรปรวนจากค่าเฉลี่ยกำลังสองนั้น ใช้สูตรเฉพาะที่สอดคล้องกับรูปแบบ

สูตรคำนวณค่าเฉลี่ยของกำลังสองของความเบี่ยงเบน (Mean Square of Deviation) หรือ MS (α) หาได้จาก

$$MS(\alpha) = SS(\alpha) / df(\alpha)$$

เมื่อ $df(\alpha)$ แทน ชั้นแห่งความเป็นอิสระที่สอดคล้องกับผลขององค์ประกอบ

$Effect(\alpha)$ แทน ผลขององค์ประกอบที่ต้องการศึกษา

$SS(\alpha)$ แทน ผลรวมกำลังสองของค่าความเบี่ยงเบน

$MS(\alpha)$ แทน ค่าเฉลี่ยกำลังสองของค่าความเบี่ยงเบน

$\sigma^2(\alpha)$ แทน ค่าประมาณความแปรปรวน

6. การวิเคราะห์ข้อมูล หลังจากผ่านขั้นตอนในการพิจารณารูปแบบการวัดและได้นำไปสู่การเก็บรวบรวมข้อมูลจากรูปแบบที่กำหนดแล้ว ข้อมูลที่ได้จะนำไปสู่การวิเคราะห์ทางสถิติเพื่อประมาณขนาดของความคลาดเคลื่อนในการวัดจากแหล่งต่างๆ กระบวนการทางสถิติที่ใช้ในทฤษฎีการสรุปอ้างอิงก็คือกระบวนการวิเคราะห์ความแปรปรวน (Analysis of Variance Procedure) โดยเฉพาะอย่างยิ่งหลังจากที่ได้เก็บรวบรวมข้อมูลแล้ว ก็จะทำการวิเคราะห์ผ่านกระบวนการ ANOVA ในขั้นตอนของการวิเคราะห์ในการศึกษา G-Study

ขั้นที่ 2 การศึกษาเพื่อการตัดสินใจ

การศึกษาเพื่อการตัดสินใจ (D-Study) เป็นขั้นตอนของการนำผลที่ได้จากการประมาณความแปรปรวนใน G-Study มีขั้นตอนดังนี้

กำหนดเอกภพของการสรุปอ้างอิงที่ผู้ตัดสินใจต้องการสรุปอ้างอิงใน D Study ซึ่งเอกภพของการสรุปอ้างอิงอาจประกอบด้วยเงื่อนไขทั้งหมดหรือเป็นเพียงเงื่อนไขย่อยในเอกภพของการสังเกตที่ยอมรับได้ ในขั้นนี้ผู้วิจัยต้องเลือกว่าจะใช้โมเดลใดเป็นโมเดลสุ่ม (random model) โมเดลกำหนด (fixed model) หรือโมเดลผสม (mixed model) ซึ่งโดยหลักการแล้วควรใช้โมเดลแบบสุ่มดีที่สุด เพราะจะทำให้การสรุป อ้างอิงได้กว้างขวางกว่าโมเดลแบบผสม และโมเดลแบบกำหนดตามลำดับ

การกำหนดขนาดตัวอย่าง (simple size) จำนวนเงื่อนไขของตัวอย่างฟาเซทใน G -Study ไม่จำเป็นต้องเป็นตัวอย่างใน D-Study ทั้งนี้การกำหนดขนาดตัวอย่างขึ้นอยู่กับความสนใจของผู้วิจัย

โครงสร้างของแบบที่ศึกษา (design structure) ใน D-Study อาจใช้โครงสร้างของแบบที่ศึกษาเหมือนหรือต่างจากใน G-Study ก็ได้ ถ้าใน D-Study ตัดสินใจให้นักเรียนทุกคนทำข้อสอบ n_i ข้อ เหมือนกันจะมีโครงสร้างเป็น D-Study ในลักษณะ pxixr design เหมือนโครงสร้างใน

G-Study ที่กล่าวมาแล้ว แต่ถ้าตัดสินใจให้นักเรียนทุกคนทำข้อสอบทุกข้อ โดยผู้ตรวจแต่ละคนทำการตรวจชุดของข้อสอบต่างกัน โดยกรณีนี้จะได้โครงสร้างเป็น p x (i:r)

การประมาณค่าความแปรปรวน (estimated D-Study variance component) ในขั้นการศึกษาเพื่อการตัดสินใจ ขึ้นอยู่กับนักวัดผลหรือนักวิจัย ดังนั้นต้องมีการประมาณค่าความแปรปรวนขึ้นมาใหม่อีกครั้ง โดยอาศัยผลจากการประมาณค่าในขั้นการศึกษาเพื่อการสรุปอ้างอิงหลักฐาน และให้สอดคล้องกับ แบบแผนและขนาดของตัวอย่างที่ต้องการตัดสินใจ

ประมาณค่าความแปรปรวนของความคลาดเคลื่อน นอกจากการประมาณค่าองค์ประกอบ ความแปรปรวน (variance component) ตามรูปแบบและขนาดของกลุ่มตัวอย่างที่ต้องการแล้วต้องประมาณค่าความแปรปรวนของความคลาดเคลื่อนแบ่งออกเป็น 2 ชนิด คือ

5.1 การประมาณค่าความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์ (Absolute Error Variance: σ^2_{Abs} หรือ σ^2_{Δ}) เป็นความแปรปรวนของผลแตกต่างระหว่างคะแนนสังเกตได้กับคะแนนเอกภพ ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่างๆ ยกเว้นความ แปรปรวนของแหล่งบุคคล ($\sigma^2_{\mu p}$ หรือ σ^2_{p}) เช่น การออกแบบ pxixr design โดยมีสูตรดังนี้

$$\sigma^2_{Abs} = \sigma^2_i + \sigma^2_r + \sigma^2_{pi} + \sigma^2_{pr} + \sigma^2_{ir} + \sigma^2_{pir,e}$$

5.2 การประมาณค่าความแปรปรวนของความคลาดเคลื่อนแบบสัมพัทธ์ (Relative Error Variance: $\sigma^2_{\mathrm{Re}l}$ หรือ σ^2_{δ}) เป็นความแปรปรวนของผลต่างระหว่างคะแนนสังเกตได้และคะแนนเอกภพ ซึ่งคำนวณได้จากผลรวมของความแปรปรวนของคะแนนจากแหล่งต่างๆ ที่มีปฏิสัมพันธ์ร่วมกับผู้สอบ (p) กับฟาเซทอื่นๆ โดยมีสูตรดังนี้

$$\sigma^2_{\mathrm{Re}l} = \sigma^2_{pi} + \sigma^2_{pr} + \sigma^2_{pir,e}$$

6. ประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เป็นขั้นตอนสุดท้ายในการศึกษา เพื่อการตัดสินใจ ซึ่งเป็นดัชนีชี้บ่งความน่าเชื่อถือของการวัด ซึ่งคำนวณได้จาก อัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพ และความแปรปรวนของคะแนนที่สังเกตที่คาดหวัง ซึ่งค่าสัมประสิทธิ์การสรุปอ้างอิง G-Coefficient สำหรับการตัดสินใจเชิงสัมพัทธ์ ($p^2_{\mathrm{re}l}$) และการตัดสินใจเชิงสัมบูรณ์ (p^2_{Abs}) สามารถคำนวณได้จากสูตรดังนี้

$$\rho^2_{\mathrm{Re}l} = \frac{\sigma^2_p}{\sigma^2_p + \sigma^2_{\mathrm{Re}l}} \quad \text{และ}$$

$$\rho^2_{Abs} = \frac{\sigma^2_p}{\sigma^2_p + \sigma^2_{Abs}}$$

ตอนที่ 5 มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบ

มโนทัศน์เบื้องต้นเกี่ยวกับทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ (Item Response Theory : IRT) เป็นทฤษฎีที่อธิบายความสัมพันธ์ระหว่างคุณลักษณะหรือความสามารถที่อยู่ภายในตัวบุคคลกับพฤติกรรมการตอบสนองข้อสอบแต่ละข้อว่ามีความน่าจะเป็นในการตอบถูกเพียงใด เป็นทฤษฎีที่ได้แก้ไขข้อจำกัดหลายๆประการของทฤษฎีการการวัดแบบดั้งเดิม (Classical Test Theory) โดยมีพัฒนาการของทฤษฎีโดยสังเขป ดังนี้ (Hambleton and Swaminathan ,1985 อ้างถึงใน เอมอร จังศิริพรปกรณ์, 2545)

ทฤษฎีนี้มีความเชื่อว่าพฤติกรรมการตอบสนองข้อสอบของผู้สอบ ซึ่งเป็นสิ่งที่สังเกตได้โดยตรงจะถูกกำหนดโดยคุณลักษณะ(trait) หรือความสามารถ (ability) ที่มีอยู่ในตัวบุคคลซึ่งไม่สามารถสังเกตได้ และยังเชื่อว่าค่าพารามิเตอร์ต่าง ๆ ของข้อสอบไม่ว่าจะเป็นค่าความยาก (b) ค่าอำนาจจำแนก (a) หรือค่าการเดา (c) ของข้อสอบแต่ละข้อเป็นคุณลักษณะที่มีอยู่ประจำ และคงที่ในตัวข้อสอบ ฉะนั้นค่าพารามิเตอร์ของข้อสอบจึงไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง และค่าความสามารถของผู้สอบซึ่งเชื่อว่าเป็นคุณลักษณะที่มีอยู่ในตัวผู้สอบนั้นจริง จึงไม่ควรแปรเปลี่ยนไปตามค่าความยากของข้อสอบ เนื่องจากความสามารถของผู้สอบเป็นคุณลักษณะซึ่งไม่สามารถวัดหรือสังเกตได้โดยตรง นักวัดผลทางการศึกษาจึงได้พยายามหาความสัมพันธ์ระหว่างปริมาณการกระทำข้อสอบ หรือคะแนน (test performance or sco re) กับปริมาณความสามารถ (ability) โดยแสดงความสัมพันธ์ดังกล่าวในลักษณะโมเดลทางคณิตศาสตร์ที่กำหนดด้วยฟังก์ชันทางคณิตศาสตร์ เรียกว่า ฟังก์ชันลักษณะข้อสอบ (Item Characteristic Function) หรือเรียกว่าโค้งคุณลักษณะของข้อสอบ (Item Characteristic Curve) หรือฟังก์ชันการตอบสนองข้อสอบ (Item Response Function) (ศิริชัย กาญจนวาสี , 2538)

สำหรับการเลือกโมเดลของทฤษฎีการตอบสนองข้อสอบมาใช้นั้น จะต้องเลือกให้เหมาะสมกับสถานการณ์ต่าง ๆ ตามข้อตกลงเบื้องต้นของแต่ละโมเดล โดยทฤษฎีนี้มีจุดเด่นหลายประการ ดังนี้ (Hambleton and Swaminathan, 1985 อ้างถึงใน กนกวรรณ รัตนธน, 2544)

ประการแรก ค่าพารามิเตอร์ของข้อสอบเป็นอิสระจากกลุ่มผู้สอบที่ใช้ในการประมาณค่า กล่าวคือ ค่าพารามิเตอร์ของข้อสอบแต่ละข้อ ได้แก่ ค่าความยาก (b) ค่าอำนาจจำแนก (a) และค่า

การเดา (c) เป็นค่าคงที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ถึงแม้ว่าผู้สอบต่างกลุ่มกัน ค่าพารามิเตอร์ของข้อสอบจะมีค่าคงเดิมเสมอ และในทำนองเดียวกันค่าพารามิเตอร์ความสามารถของผู้สอบเป็นอิสระจากข้อสอบที่ใช้ในการประมาณค่า จะไม่แปรเปลี่ยนตามค่าพารามิเตอร์ของข้อสอบ ไม่ว่าผู้สอบจะทำข้อสอบข้อใด หรือได้ทำข้อสอบเมื่อใดก็ตาม

ประการที่สอง การเปรียบเทียบความสามารถของผู้สอบ จะไม่ขึ้นอยู่กับคำถามในแบบสอบ ถึงแม้ว่าผู้สอบจะใช้ข้อคำถามต่างกัน ก็สามารถนำค่าความสามารถมาเปรียบเทียบกันได้ ทั้งนี้เนื่องจากค่าความสามารถที่ประมาณค่ามาได้นั้น เป็นคะแนนโลจิท (Logit) ซึ่งอยู่ในมาตรวัดเดียวกัน

ประการที่สาม การรายงานคุณภาพของข้อสอบในรูปค่าอินฟอร์เมชัน (information) สามารถรายงานได้ทั้งเป็นรายข้อและทั้งฉบับ ค่าอินฟอร์เมชันเป็นตัวบ่งชี้ถึงความถูกต้องแม่นยำ (accuracy) ในการประมาณค่าความสามารถ สามารถนำมาใช้แทนค่าความเที่ยง และค่าความคลาดเคลื่อนมาตรฐานในการวัดได้

ทฤษฎีการตอบสนองข้อสอบอธิบายความสัมพันธ์ระหว่างปริมาณของการกระทำหรือคะแนนของผู้สอบกับระดับความสามารถของผู้สอบ ซึ่งแสดงได้ดังสมการต่อไปนี้ (สุพัฒน์ สุกมลสันต์, 2539 อ้างถึงใน กนกวรรณ รัตนธน, 2544)

$$P = f(q)$$

เมื่อ P แทน ผลการสอบ (performance)

q แทน ความสามารถ (ability หรือ trait)

f แทน ความสัมพันธ์ (function)

จากความสัมพันธ์ในสมการดังกล่าว เป็นการแสดงความสัมพันธ์ทั่ว ๆ ไปยังไม่เฉพาะเจาะจง จึงมีผู้สนใจศึกษาและพัฒนาเพื่อให้ได้สมการที่สามารถอธิบายลักษณะเฉพาะของข้อสอบแต่ละข้อที่สามารถอธิบายความสัมพันธ์ของสองตัวแปรนี้ได้เด่นชัดขึ้น โดยอาศัยข้อตกลงเบื้องต้นบางประการและวิธีการทางคณิตศาสตร์

ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ

ทฤษฎีการตอบสนองข้อสอบ(IRT)มีข้อตกลงเบื้องต้นดังต่อไปนี้ (Hambleton and Swaminathan, 1985)

1.ความเป็นมิติเดียว (Unidimension) คุณลักษณะภายในหรือความสามารถของผู้สอบมีอยู่หลายประการ ซึ่งคุณลักษณะภายในหรือความสามารถแต่ละอย่างสามารถกำหนดพฤติกรรมการตอบสนองข้อสอบ ถ้าผลการตอบสนองข้อสอบหรือคะแนนของผู้สอบสามารถอธิบายได้คุณลักษณะเดียวก็ถือว่ามีความเป็นมิติเดียว โมเดลของการตอบสนองข้อสอบมีข้อตกลงว่า ข้อสอบแต่ละข้อในแบบสอบนั้นวัดความสามารถหรือคุณลักษณะเดียวกัน แต่ข้อตกลงข้อนี้ไม่เข้มงวดนัก ถ้ามีลักษณะเด่นที่จะวัดองค์ประกอบใดองค์ประกอบหนึ่งก็ถือว่าเป็นมิติเดียวเช่นกัน

วิธีการตรวจสอบว่าแบบสอบนั้นวัดในมิติเดียวหรือไม่นั้น มีวิธีการทดสอบได้หลายอย่าง แต่ที่นิยมมากมี 2 วิธี คือ (Warm, 1978 อ้างถึงใน สุพัฒน์ สุกมลสันต์, 2539)

1.1 โดยการหาค่าความสัมพันธ์ระหว่างค่าน้ำหนักองค์ประกอบรายข้อ (factor loading) ขององค์ประกอบที่หนึ่ง กับค่าสหสัมพันธ์แบบไบซีเรียล (biserial correlation) ของข้อสอบรายข้อกับคะแนนรวม หากค่าสหสัมพันธ์มีค่ามากกว่า .80 ก็แสดงว่าแบบสอบนั้นวัดมิติเดียว

1.2 โดยการวิเคราะห์องค์ประกอบ (factor analysis) ของข้อสอบทั้งฉบับ ถ้าค่าไอเกน (eigen value) ค่าสูงสุดแตกต่างจากค่าอื่นอย่างชัดเจนหรือไม่ ถ้าแตกต่างก็แสดงว่าแบบสอบนั้นวัดมิติเดียว

2. ความเป็นอิสระในการตอบข้อสอบ (Local Independence) หมายถึง การตอบข้อสอบข้อต่างๆ ในแบบสอบของผู้สอบนั้นมีความเป็นอิสระในเชิงสถิติ กล่าวคือ การตอบข้อสอบข้อหนึ่ง ๆ จะไม่มีผลกระทบต่อการตอบข้อสอบข้ออื่น ๆ ในแบบสอบ เนื้อหาของข้อสอบข้อหนึ่ง ๆ จะต้องไม่มีเงื่อนไขในการตอบข้อสอบข้ออื่น ๆ แมคโดแนล (McDonald, 1980a, 1980b, 1982 อ้างถึงใน Hambleton and Swaminathan ,1985) ได้นิยามของความเป็นมิติเดียวของข้อสอบว่าควรอยู่บนพื้นฐานของหลักการความเป็นอิสระในการตอบข้อสอบ

3. โค้งลักษณะของข้อสอบ (Item Characteristic Curves : ICC) เป็นฟังก์ชันทางคณิตศาสตร์ที่แสดงถึงความสัมพันธ์ระหว่างโอกาสในการตอบข้อสอบนั้นได้ถูกต้องกับระดับ

ความสามารถที่วัดได้โดยชุดของข้อสอบหรือแบบสอบ จะเห็นได้ว่าโอกาสที่ผู้สอบจะตอบข้อสอบถูก (probability) จะขึ้นอยู่กับโค้งลักษณะของข้อสอบ (ICC) ซึ่งเป็นอิสระจากการกระจายของความสามารถของผู้สอบ นั่นคือ โอกาสที่ผู้สอบจะตอบข้อสอบถูกไม่ขึ้นอยู่กับจำนวนของผู้สอบที่มีความสามารถเหมือนกัน รูปร่างของโค้งลักษณะของข้อสอบ (ICC) ในแต่ละข้อมีคุณสมบัติไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ดังนั้นจึงทำให้โอกาสในการตอบข้อสอบถูกในแต่ละข้อไม่แปรเปลี่ยน

นอกจากนี้แล้ว ถ้าข้อตกลงเบื้องต้นของการทดสอบตรงตามที่ทฤษฎีกำหนดไว้ ผลที่ได้จากการคำนวณจะมีลักษณะที่ดีดังนี้

1. ค่าพารามิเตอร์ของข้อสอบ คือ ค่าความยาก (a) ค่าอำนาจจำแนก (b) และค่าการเดา (c) จะเป็นค่าที่ไม่แปรเปลี่ยนไปตามกลุ่มผู้สอบ ไม่ว่าจะนำไปสอบกับผู้ใดก็ตาม ลักษณะเช่นนี้เรียกว่า ความเป็นอิสระจากกลุ่มตัวอย่าง (person free test calibration)

2.เมื่อทราบลักษณะการตอบข้อสอบแต่ละข้อของผู้สอบ เราสามารถจะคำนวณหาค่าความสามารถ ที่แท้จริงของผู้สอบคนนั้นได้ ค่าความสามารถที่แท้จริงนี้มีความสัมพันธ์กันโดยตรงกับคะแนนจริง การคำนวณค่าความสามารถที่แท้จริงนั้นอาจใช้ข้อสอบข้อใดก็ได้ที่วัดสิ่งเดียวกัน ลักษณะเช่นนี้เรียกว่า ความเป็นอิสระของข้อสอบ (item free person measurement)

โค้งลักษณะของข้อสอบมีหลายรูปแบบตามความเชื่อเกี่ยวกับความสัมพันธ์ระหว่างระดับความสามารถกับโอกาสในการตอบข้อสอบได้ถูกต้อง จึงทำให้มีโมเดลของโค้งลักษณะของข้อสอบขึ้นหลายโมเดล ดังจะกล่าวต่อไปนี้

โมเดลการตอบสนองข้อสอบในระยะแรกจะใช้กับแบบสอบที่ให้คะแนนรายข้อเป็น 2 ระดับ หรือให้คะแนนแบบ 0-1 ซึ่งเรียกว่าโมเดลการตอบสนองข้อสอบแบบทวิวิภาค (Dichotomous Item Response Models) แต่ถ้านำไปใช้กับข้อสอบหลายตัวเลือก ซึ่งมีการให้คะแนนบางส่วน (partial credit) หรือมาตรประมาณค่า (rating scale) จะใช้ไม่ได้ หรือจะต้องปรับการให้คะแนนเป็นแบบ 0-1 ซึ่งทำให้สูญเสียสารสนเทศของการตอบบางส่วนไป (ศิริชัย กาญจนวาสี, 2542) ต่อมาได้มีผู้พัฒนาโมเดล IRT เพื่อใช้กับแบบสอบและแบบวัดทัศนคติที่ตรวจให้คะแนนเป็นแบบมาตรประมาณค่า หรือแบบพหุวิภาค (polytomous) โมเดลในแนวนี้เรียกว่า Polytomous Item Response Models (Muraki, 1992, 1993; Donoghue, 1994; Hambleton and Zaal, 1991) การตรวจ

ให้คะแนนรายข้อมากกว่า 2 ระดับ หรือแบบพหุวิภาค สามารถจำแนกเป็นกรณีต่าง ๆ โดย ศิริชัย กาญจนวาสี (2542) ได้สรุปเป็นสาระสำคัญ ดังนี้

1. คำตอบรายข้อมีหลายรายการที่แสดงคุณลักษณะเฉพาะแตกต่างกัน(Nominal Categories) เป็น กรณีที่คำถามทุกข้อมีหลายรายการคำตอบ (Multiple-Choice Items) ซึ่งแต่ละรายการคำตอบมีคุณลักษณะเฉพาะแตกต่างกัน และรายการคำตอบไม่มีความสัมพันธ์ตามลำดับขั้นต่อกัน โมเดลที่ใช้จึงต้องทำการวิเคราะห์ทั้งคำตอบถูกและตัวเลือกทุกตัว โมเดลการตอบสนองข้อสอบที่ใช้แสดงความสัมพันธ์ของข้อมูลการตอบ ลักษณะนี้เรียกว่า Nominal Response Model (NRM) โมเดลนี้พัฒนาโดย บอค (Bock) ในปี ค.ศ. 1972

2. คำตอบรายข้อมีหลายรายการที่สามารถเรียงลำดับแต่จำนวนรายการคำตอบของแต่ละข้อไม่เท่ากัน (Specific-Category Items) เป็นกรณีที่แต่ละข้อคำถามมีจำนวนรายการหรือสเกลการตอบโดยเฉพาะ เช่น การตรวจให้คะแนนรายข้อที่มีรายการที่เป็นคำตอบถูก และรายการที่มีการให้คะแนนบางส่วน (partial credit) คะแนนเต็มของแต่ละข้ออาจไม่เท่ากัน โมเดลการตอบสนองข้อสอบนี้เรียกว่า Partial Credit Model (PCM) โมเดลนี้พัฒนาโดย มาสเตอร์ (Masters) ในปี ค.ศ. 1982

3. คำตอบรายข้อมีหลายรายการที่จัดเรียงลำดับเป็นสเกลรวมของค่าตัวเลขจำนวนเต็ม (Common Ordered Category Items) เป็นกรณีที่ทุกข้อคำถามมีรายการคำตอบร่วมกันเป็นค่าตัวเลขจำนวนเต็มเรียงตามลำดับ เช่น มาตรประมาณค่า (Rating Scale) โมเดลการตอบสนองนี้เรียกว่า Rating Scale Model (RSM) โมเดลนี้พัฒนาโดย แอนดริช (Andrich) ในปี ค.ศ. 1978

4. คำตอบรายข้อมีหลายรายการที่จัดเรียงลำดับเป็นค่าตัวเลขจำนวนเต็ม แต่จำนวนรายการคำตอบของแต่ละข้อไม่เท่ากัน (Graded Response Items) เป็นกรณีที่ทุกข้อคำถามมีรายการคำตอบเป็นค่าตัวเลขจำนวนเต็มเรียงตามลำดับ แต่จำนวนรายการของแต่ละข้อไม่จำเป็นต้องเท่ากัน โมเดลการตอบสนองนี้เรียกว่า Graded Response Model (GRM) โมเดลนี้พัฒนาโดย ซามิจิมา (Samejima) ในปี ค.ศ. 1960

จะเห็นว่าทฤษฎีการตอบสนองข้อสอบได้รับการยอมรับมากขึ้นในการทดสอบทางจิตวิทยาและทางการศึกษา เนื่องจากวิธีการสร้าง วิเคราะห์ และการให้คะแนนมีประสิทธิภาพมากกว่าทฤษฎีการทดสอบแบบดั้งเดิม (Mislery and Bock ,1990) ทฤษฎีการตอบสนองข้อสอบจึงถูกนำไปประยุกต์ใช้ในหลาย ๆ เรื่อง โดยเฉพาะการสร้างคลังข้อสอบ (Item banking) การทดสอบ

แบบปรับเหมาะกับความสามารถของผู้สอบ (Adaptive testing) การเทียบคะแนนการสอบ (Test score equating) การหาความลำเอียงของข้อสอบ (Item bias)

สำหรับรายละเอียดของโมเดลการตอบสนองข้อสอบแบบทวิวิภาค และโมเดลการตอบสนองข้อสอบแบบพหุวิภาค จะนำเสนอในหัวข้อต่อไป

มโนทัศน์เกี่ยวกับทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค
(Polytomous Item Response Theory)

โมเดล IRT ที่พัฒนาในระยะแรกจะใช้กับแบบสอบที่ให้คะแนนแบบทวิวิภาค (dichotomous) เป็นส่วนใหญ่ ต่อมาได้มีผู้พัฒนาโมเดล IRT เพื่อใช้กับแบบสอบและแบบวัดทัศนคติที่ตรวจให้คะแนนแบบพหุวิภาค (polytomous) หรือลักษณะมาตรประมาณค่า (rating data) โมเดลในแนวทฤษฎีนี้ทั้งหมดเรียกว่า (Polytomous Item Response Models (Muraki, 1992, 1993; Donoghue, 1994; Hambleton and Zaal, 1991) มีโมเดลที่สำคัญ คือ

1. Graded Response Model (GRM) พัฒนาโดย ซามิจิมา (Samejima) ในปี ค.ศ. 1960
2. Normal Response Model (NRM) พัฒนาโดย บอค (Bock) ในปี ค.ศ. 1972
3. Continuous Model (CM) พัฒนาโดย ซามิจิมา ในปี ค.ศ.1973
4. Rating Scale Model (RSM) พัฒนาโดย แอนดริช (Andrich) ในปี ค.ศ. 1978
5. Partial Credit Model (PCM) พัฒนาโดย มาสเตอร์ (Masters) ในปี ค.ศ. 1982
6. Sucessive Interval Model (SIM) พัฒนาโดย รอส์ท (Rost) ในปี ค.ศ. 1988
7. Generalized Partial Credit Model : GPCM พัฒนาโดย มูราคิ (Muraki) ในปี ค.ศ. 1992

โมเดลที่มีผู้ศึกษากันมากในปัจจุบัน ไม่เข้มงวดเกี่ยวกับข้อตกลงเบื้องต้น และสามารถใช้ได้กับแบบสอบและแบบวัดหลายลักษณะคือ GRM และ GPCM (Donoghue, 1994; De Ayala, 1994; Muraki, 1992, 1993; Reise and Yu, 1990; Koch and De Ayala, 1989; Koch, 1983)

ซามิจิมา (Samejima) ได้พัฒนา GRM จากโมเดลโลจิสติก 2 พารามิเตอร์ เพื่อใช้กับแบบสอบและแบบวัดทัศนคติที่ตรวจให้คะแนนแบบพหุวิภาค (polytomous) ในขณะที่มูราคิ (Muraki) ได้พัฒนา GPCM จาก PCM ของ ไรท์ และมาสเตอร์ (Wright and Masters, 1982) ให้สามารถใช้ได้กับแบบสอบและแบบวัดที่มีลักษณะเป็นพหุมิติ โดยรวมค่าอำนาจจำแนกข้อกระทงในการประมาณค่าพารามิเตอร์ของฟังก์ชันด้วยโมเดลทั้ง 2 เมื่อพิจารณาดูแล้วคล้ายคลึงกันมาก เพียงแต่มีมโนทัศน์

เกี่ยวกับฟังก์ชันลักษณะเชิงปฏิบัติ (operating characteristic function : OCF) และพัฒนาการที่แตกต่างกัน รูปแบบทั่วไปของโมเดลในแนวคิดนี้พัฒนามาจากสูตร

เมื่อ $P_j(U_j=1/q)$ คือ ความน่าจะเป็นของคนที่ i มีความสามารถ สามารถตอบข้อสอบข้อ j (U=1) ได้ถูกต้อง

a_j คือ ค่าพารามิเตอร์อำนาจจำแนก

b_j คือ ค่าพารามิเตอร์ของความยากของข้อสอบ

exp คือ ค่าคงที่ 2.7183

$$P_i(U_j = 1/\theta) = \frac{\exp^{[a_j(\theta - b_j)]}}{1+\exp^{[a_j(\theta - b_j)]}}$$

ดังนั้น ในการศึกษาครั้งนี้จึงขอนำเสนอเฉพาะโมเดล GRM ดังมีรายละเอียดพอสังเขป ดังนี้

GRM (Graded Response Model)

GRM มีข้อตกลงว่าการตอบสนองข้อกระทง j ของผู้ตอบ สามารถแบ่งคะแนนออกเป็น mj+1 ลำดับขั้น (categories) คะแนนแต่ละลำดับ (j_k) กำหนดให้เป็น 0, 1, 2, …, mj ตามลำดับ (Donoghue, 1994; Koch, 1983) เขียนสมการทั่วไปได้ดังนี้

$$P_{jk}(\theta) = \frac{\exp^{[Da_j(\theta - b_{jk})]}}{1+\exp^{[Da_j(\theta - b_{jk})]}}$$

เมื่อ $P_{jk}(q)$ คือ ความน่าจะเป็นที่ผู้มีคุณลักษณะ q จะตอบข้อกระทง j ได้คะแนน k

D คือ ค่าคงที่ของสเกล (scaling constant) เมื่อปรับโค้งฟังก์ชันของโมเดลโลจิสติก และโมเดลนอร์มอลออกไจฟ์ มีค่าเท่ากับ 1.70

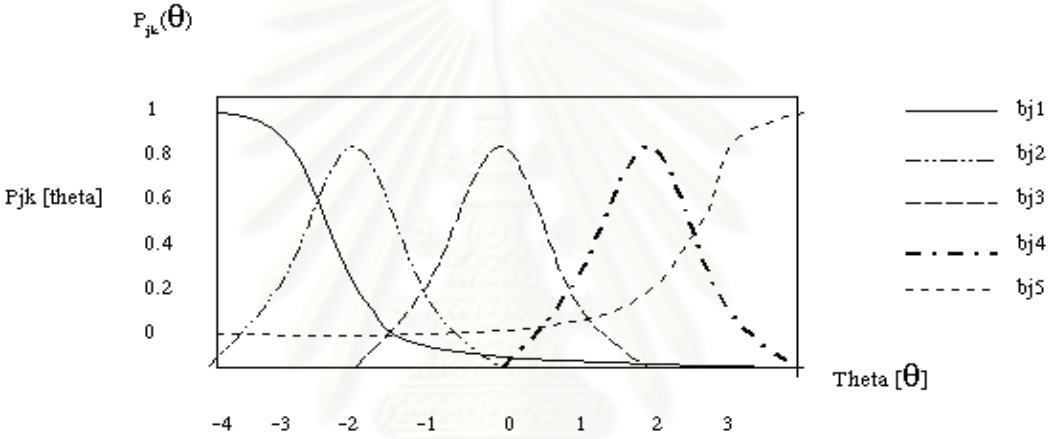
a_j คือ ค่าอำนาจจำแนกของข้อความ j

q คือ ความสามารถหรือคุณลักษณะภายในของผู้สอบ

b_{jk} คือ ค่าความยากของข้อความ j ในลำดับขั้นที่ jk เมื่อ jk = 0, 1, …, k

exp คือ ค่าคงที่ มีค่าเท่ากับ 2.7183

ฟังก์ชันทางคณิตศาสตร์ดังกล่าวสามารถเขียนแทนด้วยโค้งฟังก์ชันสารสนเทศได้ดังแผนภาพที่ 2.7 (Koch, 1983)



แผนภาพที่ 2.7 โค้งฟังก์ชันสารสนเทศของ GRM

ในโมเดลนี้แสดงถึงความสัมพันธ์ของคนที่มีคุณลักษณะภายใน (q) สูง มีความน่าจะเป็นในการตอบลำดับขั้นคะแนนที่สูงกว่าคนที่มีคุณลักษณะภายในต่ำ ซึ่งนักวัดผลได้มีการประยุกต์ใช้โมเดลนี้กับแบบสอบชนิดปรับเหมาะ (Adaptive Testing) และแบบวัดทัศนคติต่าง ๆ

ลักษณะเฉพาะฟังก์ชันของโมเดล GRM สรุปได้ดังนี้

1.1) ฟังก์ชันสารสนเทศของแบบวัด (TIF) เขียนแทนด้วยสมการดังนี้

$$I(\theta) = \sum_{j=1}^{m} I_j(\theta)$$

เมื่อ I(q) คือ ฟังก์ชันสารสนเทศของข้อกระทงตั้งแต่ข้อ 1 ถึงข้อที่ m

1.2) ฟังก์ชันสารสนเทศของข้อกระทง (IIF) สามารถแทนด้วยสมการดังนี้

$$I_j(\theta) = \sum_{jk=0}^{k} \frac{[P'_{jk}(\theta)]^2}{P_{jk}(\theta)}$$

เมื่อ $P'_{jk}(q)$ คือ ความน่าจะเป็นสำหรับคนที่มีความสามารถ q จะตอบข้อกระทง j ได้คะแนน 1 มากกว่าคะแนน k ใด ๆ

$P_{jk}(q)$ คือ ความน่าจะเป็นสำหรับคนที่มีความสามารถ q จะตอบข้อกระทง j ได้คะแนน k

1.3) ฟังก์ชันสารสนเทศของแต่ละลำดับขั้นคะแนน (ICIFs) เขียนแทนด้วยสมการดังนี้

$$I_j(\theta) = \frac{[P'_{jk}(\theta)]^2}{P_{jk}(\theta)}$$

เมื่อ $P'_{jk}(q)$ คือ ความน่าจะเป็นสำหรับคนที่มีความสามารถ q จะตอบข้อกระทง j ในลำดับขั้นคะแนนที่ 1 ได้ถูกต้องมากกว่าคะแนน k ใด ๆ

$P_{jk}(q)$ คือ ความน่าจะเป็นสำหรับคนที่มีความสามารถ q จะตอบข้อกระทง j ในลำดับขั้นคะแนน k ได้ถูกต้อง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 6 งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยที่เกี่ยวข้อง ผู้วิจัยขอนำเสนอเป็น 3 ตอน คือ งานวิจัยที่เกี่ยวข้องกับแผนผังมโนทัศน์ งานวิจัยที่เกี่ยวข้องกับทฤษฎีสรุปอ้างอิง และงานวิจัยที่เกี่ยวข้องกับทฤษฎีตอบสนองข้อสอบ ซึ่งมีรายละเอียดดังนี้

งานวิจัยที่เกี่ยวข้องกับแผนผังมโนทัศน์

จากการศึกษางานวิจัยที่เกี่ยวข้องกับแผนผังมโนทัศน์ นักการศึกษาส่วนใหญ่ในต่างประเทศได้ศึกษาการใช้แผนผังมโนทัศน์ในการประเมิน ส่วนในประเทศไทยส่วนใหญ่ใช้ในการจัดกิจกรรมการเรียนการสอน มีเพียงบางงานวิจัยที่นำมาใช้ในการประเมิน ซึ่งสรุปได้ดังนี้

ต่างประเทศ

บอลท์ (Bolte, 1999) ได้ศึกษาการใช้แผนผังมโนทัศน์และการเขียนอธิบายในแผนผังมโนทัศน์ประเมินผลในวิชาคณิตศาสตร์ โดยกลุ่มตัวอย่างคือ นักศึกษามหาวิทยาลัยในภาคตะวันตกเฉียงเหนือของสหรัฐอเมริกา จำนวน 3 กลุ่ม โดยแต่ละกลุ่มได้รับการสอนวิธีสร้างแผนผังมโนทัศน์แล้วประเมินผลการสร้างแผนผังมโนทัศน์ของนักศึกษาและให้นักศึกษาเขียนคำอธิบายเนื้อหาในแผนผังมโนทัศน์ด้วย การศึกษาครั้งนี้ดูจากผลของแผนผังมโนทัศน์ การออกแบบการวิจัย การประเมินเกี่ยวกับความรู้วิชาคณิตศาสตร์ ความสัมพันธ์ของวิธีประเมินแบบประเพณีนิยม แผนผังมโนทัศน์และการเขียนอธิบายเป็นการช่วยเพิ่มความรู้ของนักเรียน ผลการวิจัยพบว่าแผนผังมโนทัศน์และการเขียนอธิบายความในแผนผังมโนทัศน์ใช้ในการประเมินผลได้ดีและสามารถทำให้วิธีการช่วยสร้างองค์ความรู้ในวิชาคณิตศาสตร์ให้กับผู้เรียน และนอกจากนี้การใช้แผนผังมโนทัศน์และการเขียนอธิบายความเรียงในแผนผังมโนทัศน์ในการประเมินผลมีความสัมพันธ์กับคะแนนจากการทดสอบในชั้นเรียน (course exam) และผลการเรียน (final grade) อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

แมคเคลอร์ และคณะ(McClure et al, 1999) ได้ศึกษาถึงการนำแผนผังมโนทัศน์มาใช้ในการประเมินในชั้นเรียน ด้านความเที่ยง ความตรง และความเป็นไปได้ในทางปฏิบัติ ได้ทำการศึกษาวิจัยเกี่ยวกับนักศึกษาที่เพิ่งเข้ามาใหม่ในปี 1 ที่ลงทะเบียนเรียนวิชาทางจิตวิทยา โดยแบ่งเป็น 2 กลุ่ม ผู้วิจัยได้ให้ความรู้แก่กลุ่มตัวอย่างเกี่ยวกับเทคนิคแผนผังมโนทัศน์ กลุ่มละ 90 นาที ผลการศึกษาพบว่าด้านความเที่ยงใช้เปรียบเทียบวิธีการให้คะแนน 6 วิธี พบว่า มีค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง .23 - .76 วิธีการตรวจให้คะแนนที่ให้ความเที่ยงสูงที่สุดคือ วิธีการตรวจให้คะแนนแผนผังโดยดูจากการเชื่อมมโนทัศน์และคำเชื่อม ส่วนในด้านความตรงตามสภาพพบว่า

แผนผังมโนทัศน์ที่นักเรียนสร้างขึ้นสอดคล้องกับแผนผังมโนทัศน์เกณฑ์ของผู้วิจัย และมีความเป็นไปได้ที่จะนำแผนผังมโนทัศน์มาใช้ในการประเมินผลในห้องเรียน

รุซ-พรีโม และคณะ (Ruiz-Primo et al, 2001) ได้ศึกษาเปรียบเทียบเทคนิคแผนผังมโนทัศน์ 2 แบบ โดยรูปแบบแผนผังมโนทัศน์กำหนดทิศทางให้สูง คือ แผนผังมโนทัศน์แบบเติมคำลงในผัง และแผนผังแบบกำหนดทิศทางให้ต่ำ คือแผนผังมโนทัศน์ที่สร้างผังจากคำ ซึ่งผู้วิจัยทำการทดสอบคุณภาพในด้าน 1) คะแนนของแผนผังมโนทัศน์โครงสร้างมีความไวต่อการเลือกเติมคำมโนทัศน์หรือเติมคำบนเส้นเชื่อม 2) แผนผังโครงสร้างของทั้งสองแบบมีลักษณะคู่ขนานกัน 3) เทคนิคแผนผังทั้งสองแบบให้ข้อมูลสารสนเทศเกี่ยวกับความเข้าใจในการเชื่อมโยงของนักเรียนคล้ายๆกัน กลุ่มตัวอย่างที่ใช้ในงานวิจัยเป็นนักเรียนระดับมัธยมศึกษา จำนวน 152 คน และครูผู้สอนวิชาเคมี จำนวน 2 คน โดยกลุ่มตัวอย่างถูกฝึกให้สร้างแผนผังมโนทัศน์ทั้งสองรูปแบบโดยใช้เวลา 50 นาที ผลการวิจัยพบว่านักเรียนไม่มีความไวต่อการเลือกการเติมคำมโนทัศน์หรือคำเชื่อมบนเส้น ถึงแม้ว่าเทคนิคแผนผังมโนทัศน์แบบเติมคำลงในผังและการเติมคำเชื่อมบนเส้นจะมีลักษณะไม่เป็นคู่ขนานกันกับรูปแบบการเติมคำในแผนผัง และแผนผังมโนทัศน์ที่มีการกำหนดทิศทางสูงและแผนผังที่กำหนดทิศทางต่ำมีความแตกต่างกันเกี่ยวกับการอธิบายองค์ความรู้ของนักเรียน อย่างไรก็ตามคะแนนที่ได้รับจากเทคนิคที่มีการกำหนดทิศทางให้ต่ำ พบว่าองค์ความรู้ของนักเรียนไม่สมบูรณ์เมื่อเทียบกับแผนผังเกณฑ์ สรุปได้ว่าเทคนิคแผนผังมโนทัศน์แบบสร้างผังจากคำส่งผลต่อความสร้างองค์ความรู้ของนักเรียนได้ดีกว่าแผนผังมโนทัศน์แบบเติมคำลงในผัง

เคน และคณะ (Klein et al, 2002) ได้ทำการทดสอบความตรงของแผนผังมโนทัศน์ในการใช้เป็นเครื่องมือสำหรับการประเมินในวิชาวิทยาศาสตร์ ในการศึกษาครั้งนี้ได้แบ่งออกเป็น 3 ด้าน คือ 1) ให้เค้าโครงของโมเดลในการสร้างงานแผนผังมโนทัศน์, เสนอชุดมาตรฐานของขั้นตอนและการใช้เนื้อหาของการคิด 2) อธิบายวิธีการให้คะแนนที่ใช้ในการประเมินการปฏิบัติงานของนักเรียน รวมทั้งการอภิปรายถึงความเที่ยงของวิธีการและความสัมพันธ์ของวิธีการนั้นกับวิธีการให้คะแนนอย่างอื่นที่เป็นไปได้ 3) เสนอผลการวิเคราะห์ทางสถิติรวมทั้งการวิเคราะห์เปรียบเทียบ มีการวิเคราะห์ความตรงเชิงโครงสร้างใน 2 ลักษณะและวิธีการวัดที่แตกต่างกัน 3 วิธี การวิเคราะห์ประพจน์ และการวิเคราะห์การให้รายละเอียดของประพจน์ ผลการวิจัยแสดงให้เห็นว่า แผนผังมโนทัศน์มีความไวต่อระดับความสามารถของนักเรียนโดยผลจากวิธีพหุลักษณะพหุวิธี

หยิน และคณะ (Yin et al, 2004) ทำการทดสอบความเท่าเทียมกันของเทคนิคแผนผังมโนทัศน์ 2 แบบ 1) เทคนิคแผนผังมโนทัศน์แบบสร้างผังจากคำโดยสร้างคำเชื่อมเอง--C และ 2) เทคนิคแผนผังมโนทัศน์แบบสร้างผังจากคำโดยการเลือกคำเชื่อมจากที่กำหนดมาให้--S โดยได้เปรียบเทียบ 2 ลักษณะคือผลลัพธ์และกระบวนการของแผนผังมโนทัศน์ ซึ่งการศึกษานี้

ประกอบด้วยตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพ ผลการศึกษาพบว่า แผนผังมโนทัศน์แบบสร้างผังจากคำโดยสร้างคำเชื่อมเอง--C และแผนผังมโนทัศน์แบบสร้างผังจากคำโดยการเลือกคำเชื่อมจากที่กำหนดมาให้--S ไม่เท่าเทียมกันโดยที่ C ดีกว่า S ในการที่สามารถตรวจสอบความรู้บางส่วนของนักเรียนได้ แต่ S จะสามารถตรวจให้คะแนนได้ง่ายและมีประสิทธิภาพมากกว่า C และจากคุณลักษณะทั้งหมดของเทคนิคแผนผังมโนทัศน์ทั้ง 2 เทคนิค ถ้านำไปใช้ในการประเมินผลแล้ว C เหมาะสมที่จะทำไปใช้ในการประเมินผลระหว่างเรียน และ S นำไปใช้ได้ดีสำหรับการประเมินผลสำคัญๆ ขนาดใหญ่ (large-scale assessment)

ในประเทศ

รุ่งนภา ทศภานนท์ (2544) ได้ทำการศึกษาเรื่องผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนโดยใช้เทคนิคการจัดแผนผังมโนทัศน์ในการเรียนการสอนวิชาคณิตศาสตร์ ,เปรียบเทียบผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียน ชั้นมัธยมศึกษาปีที่ 4 ระหว่างกลุ่มที่เรียนโดยใช้เทคนิคการจัดแผนผังมโนทัศน์ในการเรียน การสอนวิชาคณิตศาสตร์และกลุ่มที่มีการเรียนการสอนแบบปกติ ,ศึกษาความสามารถในการสร้างแผนผังมโนทัศน์ทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนโดยใช้เทคนิค การจัดแผนผังมโนทัศน์ในการเรียนการสอนวิชาคณิตศาสตร์ และศึกษาความสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์และความสามารถในการสร้างแผนผังมโนทัศน์ทางคณิตศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนโดยใช้เทคนิคการจัดแผนผังมโนทัศน์ในการเรียน การสอนวิชาคณิตศาสตร์ ตัวอย่างประชากรเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนตระกูลประเทืองวิทยาคม จังหวัดยโสธร ปีการศึกษา 2544 จำนวน 66 คน โดยแบ่งเป็นกลุ่มทดลองที่เรียนโดยใช้ เทคนิคการจัดแผนผังมโนทัศน์ในการเรียนการสอนวิชาคณิตศาสตร์และกลุ่มควบคุมที่มีการเรียนการสอนแบบปกติ เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ แบบทดสอบวัดผลสัมฤทธิ์ทาง การเรียนคณิตศาสตร์ ที่มีค่าความเที่ยงเท่ากับ 0.91 และแบบทดสอบวัดความสามารถในการสร้างแผนผังมโนทัศน์ วิเคราะห์ข้อมูลโดยการหาค่าเฉลี่ยเลขคณิตคิดเป็นร้อยละ มัธยฐาน คิดเป็นร้อยละ ฐานนิยมคิดเป็นร้อยละ ความเบ้ (Skewness) ความโด่ง (Kurtosis) ค่าที (t-test) และการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันโพรดัคโมเมนต์ ผลการวิจัยพบว่า 1)นักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนโดยใช้เทคนิคการจัดแผนผังมโนทัศน์ในการเรียนการสอนวิชาคณิตศาสตร์มีผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ผ่านเกณฑ์ร้อยละ 50 ที่กำหนดไว้ 2) นักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนโดยใช้เทคนิคการจัดแผนผังมโนทัศน์ในการเรียนการสอนวิชาคณิตศาสตร์มีผลสัมฤทธิ์ทางการเรียนสูงกว่านักเรียนกลุ่มที่เรียนแบบปกติ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 3) นักเรียนชั้นมัธยมศึกษาปีที่ 4 ที่เรียนโดยใช้

เทคนิคการจัดแผนผังมโนทัศน์ในการ เรียนการสอนวิชาคณิตศาสตร์มีความสามารถในการสร้างแผนผังมโนทัศน์ทางคณิตศาสตร์ผ่าน เกณฑ์ร้อยละ 50 ที่กำหนดไว้ 4) ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์มีความสัมพันธ์ทางบวกกับความสามารถในการ สร้างแผนผังมโนทัศน์ทางคณิตศาสตร์

วิยะดา ระวังสุข (2545) ได้พัฒนาแบบประเมินความคิดรวบยอดวิชาวิทยาศาสตร์โดยใช้แผนผังมโนทัศน์ และได้ศึกษาคุณภาพของแบบประเมินความคิดรวบยอดวิชาวิทยาศาสตร์ ในด้าน 1) คุณภาพของแผนผังมโนทัศน์แบบเติมคำลงในผังด้านความเที่ยงแบบสอดคล้องภายใน โดยใช้สูตรสัมประสิทธิ์แอลฟาของครอนบาค 2) เปรียบเทียบค่าสัมประสิทธิ์สรุปอ้างอิงของคะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำ เมื่อตรวจให้คะแนนต่างประเภทกัน 3) ศึกษาความตรงตามสภาพโดยพิจารณาจากความสอดคล้องระหว่างผลการประเมินโดยใช้แผนผังมโนทัศน์กับผลการวินิจฉัยนักเรียนเป็นรายบุคคลโดยครูผู้สอน และศึกษาความเหมาะสมและความเป็นไปได้ในการนำแผนผังมโนทัศน์ไปใช้ในการประเมิน กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 4 โรงเรียนวัดลำนาว จังหวัดนครศรีธรรมราช ปีการศึกษา 2545 จำนวน 38 คน เครื่องมือที่ใช้ในการวิจัยได้แก่ แผนผังมโนทัศน์ การวิเคราะห์ข้อมูลโดยการหาค่าสถิติพื้นฐาน การประมาณค่าสัมประสิทธิ์สรุปอ้างอิง ผลการวิจัยพบว่าแบบประเมินความคิดรวบยอดที่พัฒนาขึ้นเป็นแผนผังมโนทัศน์แบบเติมคำและแผนผังมโนทัศน์แบบสร้างผังจากคำ โดยมีเนื้อหาสาระในวิชาวิทยาศาสตร์ ซึ่งแผนผังมโนทัศน์แบบเติมคำลงในผังมีค่าความเที่ยงในสาระที่ 1 และที่ 2 เท่ากับ .46 และ .76 ตามลำดับ และจากการเปรียบเทียบการตรวจให้คะแนน 3 แบบ พบว่า การตรวจแบบ convergence score มีค่าสัมประสิทธิ์สรุปอ้างอิงสูงสุด ส่วนในด้านความตรงตามสภาพพบว่ามีค่าความสอดคล้องอยู่ระหว่าง.72-.92 และ .40-.68 และจากการสัมภาษณ์ครูและนักเรียน พบว่าแผนผังมโนทัศน์ที่พัฒนาขึ้นมีความเหมาะสมและความเป็นไปได้ในการนำไปใช้ประเมินในชั้นเรียน

ณัฐยากร แกล้วทนงค์ (ณัฐยากร แกล้วทนงค์, 2545) ได้ทำการศึกษาผลของการใช้เทคนิคแผนผังความคิดต่อผลสัมฤทธิ์ทางการอ่านของนักเรียนที่มีระดับความสามารถทางภาษาต่างกัน กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 5 โรงเรียนบ้านชะอวด จังหวัดนครศรีธรรมราช จำนวน 120 คน ผลการวิจัยพบว่าการอ่านโดยใช้เทคนิคแผนผังความคิดส่งผลต่อผลสัมฤทธิ์ทางการอ่านสูงกว่าการอ่านแบบปกติอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 นักเรียนที่มีความสามารถทางภาษาสูงมีผลสัมฤทธิ์ทางการอ่านสูงกว่านักเรียนที่มีระดับความสามารถทางภาษา

ต่ำอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และไม่มีปฏิกิริยาร่วมระหว่างวิธีการอ่านและระดับความสามารถทางภาษา

กัญญา ทิพย์ลาย (2545) ได้ศึกษาเกี่ยวกับการสร้างแบบฝึกการเขียนสรุปความจากร้อยแก้วสำหรับนักเรียนชั้นประถมศึกษาปีที่ 5 โดยใช้แผนผังความคิด เพื่อหาประสิทธิภาพตามเกณฑ์มาตรฐาน 80/80 และเพื่อเปรียบเทียบผลสัมฤทธิ์การเขียนสรุปความระหว่างคะแนนก่อนใช้แบบฝึกและหลังจากใช้แบบฝึก กลุ่มตัวอย่างที่ใช้ในการศึกษาคือนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนวัดอัมพวนาราม จังหวัดพัทลุง จำนวน 25 คน ผลการวิจัยพบว่าแบบฝึกการเขียนสรุปความจากบทร้อยแก้ว สำหรับนักเรียนชั้นประถมศึกษาปีที่ 5 โดยใช้แผนผังความคิดที่ผู้วิจัยสร้างขึ้นมีประสิทธิภาพ 81.07/80.56 และผลสัมฤทธิ์ทางการเขียนสรุปความหลังใช้แบบฝึกสูงกว่าก่อนใช้แบบฝึกอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

ธัญพันธ์ พงศาวกุล (2545) ได้ทำการศึกษาเพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเขียนเรียงความของนักเรียนก่อนกับหลังได้รับการสอนโดยใช้แผนผังมโนทัศน์ และเพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเขียนเรียงความของนักเรียนก่อนกับหลังได้รับการสอนแบบปกติ และเพื่อเปรียบเทียบผลสัมฤทธิ์ทางการเขียนของนักเรียนที่ได้รับการสอนโดยใช้แผนผังมโนทัศน์กับการสอนแบบปกติ กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2545 โรงเรียนมุจลินทวาปีวิหาร (เพชรานุกูลกิจ) จังหวัดปัตตานี จำนวน 80 คน ผลการวิจัยพบว่านักเรียนที่ได้รับการสอนโดยใช้แผนผังมโนทัศน์มีผลสัมฤทธิ์ทางการเขียนเรียงความสูงกว่าก่อนได้รับการสอน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 นักเรียนที่ได้รับการสอนแบบปกติมีผลสัมฤทธิ์ทางการเขียนเรียงความสูงกว่าก่อนได้รับการสอน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001 และนักเรียนที่ได้รับการสอนโดยใช้แผนผังมโนทัศน์มีผลสัมฤทธิ์ทางการเขียนเรียงความสูงกว่านักเรียนที่ได้รับการสอนแบบปกติ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.001

สุทธิพร วงศ์วิชัย (2545) ได้ศึกษาความคิดรวบยอดและวิธีคิดในการเรียนวิชาวรรณคดีของนักเรียนชั้นมัธยมศึกษาปีที่ 5 ที่เรียนด้วยการใช้เทคนิคกลุ่มร่วมมือโดยเน้นการสร้างแผนภูมิมโนทัศน์ กลุ่มตัวอย่างคือ นักเรียนชั้นมัธยมศึกษาปีที่ 5/3 โรงเรียนสาธิตมหาวิทยาลัย เชียงใหม่จำนวน 39 คน ผลการวิจัยพบว่า ความคิดรวบยอดในการเรียนวรรณคดีไทย หลังการเรียนด้วยการใช้เทคนิคกลุ่มร่วมมือโดยเน้นการสร้างแผนภูมิมโนทัศน์สูงกว่าก่อนเรียน ในส่วนค่าเฉลี่ยของคะแนนที่ได้จากแบบฝึกการสร้างแผนภูมิมโนทัศน์ขณะทำการสอน จำนวน 4 ครั้ง พบว่ามีแนวโน้มสูงขึ้นจากครั้งที่ 1 ถึงครั้งที่ 4 ส่วนแบบทดสอบการสร้างแผนภูมิมโนทัศน์แบบอัตนัยโดยใช้ทดสอบเฉพาะหลังเรียน เพื่อหาวิธีคิดของนักเรียน พบว่านักเรียนมีพัฒนาการการสร้างความคิดรวบยอดในการเรียนวรรณคดีไทยโดยการใช้การสร้างแผนภูมิมโนทัศน์

ประกิต ผลมูล (2546) ศึกษาแผนภูมิมโนมิติ มโนมิติที่คลาดเคลื่อน การใช้แผนภูมิมโนมิติเพื่อปรับเปลี่ยนมโนมิติที่คลาดเคลื่อนในวิชาชีววิทยาเรื่องรากและลำต้น ของนักเรียนให้มีมโนมิติที่ถูกต้องและมีมโนมิติที่คลาดเคลื่อนน้อยลง กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 5 จำนวน 32 คน ผลการวิจัยพบว่า หลังจากที่นักเรียนที่เป็นกลุ่มตัวอย่างผ่านการทดลองแล้วมีคะแนนรวมของแผนภูมิสุดท้าย (ครั้งที่ 4) แตกต่างจากแผนภูมิมโนมิติย่อยครั้งที่ 1,2 และ 3 อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 นักเรียนมีการปรับเปลี่ยนมโนมิติที่คลาดเคลื่อนเป็นมโนมิติที่ถูกต้อง 10 ข้อจากข้อสอบวัดมโนมิติที่คลาดเคลื่อนทั้งหมด 17 ข้อ นักเรียนมีมโนมิติที่คลาดเคลื่อนหลังเรียนลดน้อยลงกว่าก่อนเรียน โดยมีจำนวนนักเรียนที่สอบได้คะแนนผ่านเกณฑ์คะแนนจุดตัดมากกว่าก่อนเรียนและมีคะแนนการทำแบบทดสอบวัดมโนมิติที่คลาดเคลื่อนหลังเรียนสูงกว่าก่อนเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ตารางที่ 2.4 สรุปผลการศึกษาในต่างประเทศและในประเทศเกี่ยวกับเทคนิคการใช้แผนผังมโนทัศน์และการให้คะแนนแผนผังมโนทัศน์

แหล่งอ้างอิง	เนื้อหา (ขนาดกลุ่มตัวอย่าง)	เทคนิคแผนผังมโนทัศน์	ระบบการให้คะแนน
McMclure et al (1999)	Educational Psychology (n=63)	Construct a map	1. Holistic 2. Holistic with master map 3. Relational 4. Relational with master map 5. Structural 6. Structural with master map
Bolte (1999)	Mathematics (n=103)	Construct a map	1. Organization 2. Accuracy
Ruiz-Primo et al (2001)	Science (n=152)	Construct a map Fill in the map	1. Proposition accuracy score 2. Convergence score 3. Salience score
Klein et al (2002)	Hearing and Vision (n=56)	Construct a map	1. Original 2. Supplementary credit 3. Quantity Rating
Yin et al (2004)	Science (n=94)	Construct a map - created linking phrase - selected linking phrase	1. Proposition accuracy score

ตารางที่ 2.4 สรุปผลการศึกษาในต่างประเทศและในประเทศเกี่ยวกับเทคนิคการใช้แผนผังมโนทัศน์และการให้คะแนนแผนผังมโนทัศน์ (ต่อ)

แหล่งอ้างอิง	เนื้อหา (ขนาดกลุ่มตัวอย่าง)	เทคนิคแผนผังมโนทัศน์	ระบบการให้คะแนน
รุ่งนภา ทศภานนท์ (2544)	คณิตศาสตร์ (n=66)	Fill-in task Add-on task Close-ended list task Open-ended list task	1. Organization 2. Accuracy
วิยะดา ระวังสุข (2545)	วิทยาศาสตร์ (n=38)	Construct a map	1. Original 2. Supplementary credit 3. Quantity Rating
ณัฐยากร แกล้วทนงค์ (2545)	การอ่าน (n=120)	Construct a map	No information provided
กัญญา ทิพย์ลาย (2545)	การสรุปความ (n=80)	Construct a map	No information provided
ธัญพันธ์ พงศาวกุล (2545)	การเขียนเรียงความ (n=39)	Construct a map	No information provided
สุทธิพร วงศ์วิชัย (2545)	วรรณคดี (n=39)	Construct a map	No information provided
ประกิต ผลมูล (2546)	ชีววิทยา (n=32)	Construct a map	1. Organization 2. Accuracy

จากตารางที่ 2.4 สรุปผลการศึกษาในต่างประเทศและในประเทศเกี่ยวกับเทคนิคการใช้แผนผังมโนทัศน์และการให้คะแนนแผนผังมโนทัศน์ พบว่างานวิจัยส่วนใหญ่นำเทคนิคแผนผังมโนทัศน์แบบสร้างผังจากคำ (Construct a map) มาใช้ในการศึกษาวิจัย

ตารางที่ 2.5 สรุปผลการนำแผนผังมโนทัศน์ไปใช้

แหล่งอ้างอิง	เทคนิคแผนผังมโนทัศน์	การนำไปใช้	
		การเรียนการสอน	การประเมินผล
McMclure et al (1999)	Construct a map		✓
Bolte (1999)	Construct a map		✓
Ruiz-Primo et al (2001)	Construct a map Fill in the map		✓
Klein et al (2002)	Construct a map		✓
Yin et al (2004)	Construct a map - created linking phrasse - selected linking phrase		✓
รุ่งนภา ทศภานนท์ (2544)	Fill-in task Add-on task Close-ended list task Open-ended list task	✓	✓
วิยะดา ระวังสุข (2545)	Construct a map		✓
ณัฐยากร แกล้วทนงค์ (2545)	Construct a map	✓	
กัญญา ทิพย์ลาย (2545)	Construct a map	✓	
ธัญพันธ์ พงศาวกุล (2545)	Construct a map	✓	
สุทธิพร วงศ์วิชัย (2545)	Construct a map	✓	
ประกิต ผลมูล (2546)	Construct a map		✓

จากตารางที่ 2.5 สรุปผลการนำแผนผังมโนทัศน์ไปใช้ได้ว่างานวิจัยส่วนใหญ่ในประเทศนำไปศึกษาในด้านการเรียนการสอน ในขณะที่งานวิจัยต่างประเทศนั้นนำไปศึกษาด้านการประเมินผลมากกว่า

งานวิจัยที่เกี่ยวข้องกับทฤษฎีสรุปอ้างอิง

สุพิมพ์ ศรีพันวรสกุล (2538) ทำการวิเคราะห์ความตรงลู่เข้าของการวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ตามทฤษฎีการอ้างอิงสรุป ด้วยการวิเคราะห์และเปรียบเทียบความตรงลู่เข้าจากสูตรที่พัฒนาตามทฤษฎีนี้ในองค์ประกอบของข้อกระทงและวิธีวัด และทำการตรวจสอบความตรงข้ามกลุ่มจากกลุ่มตัวอย่างที่ใช้ศึกษาและกลุ่มตัวอย่างสอบทานผล ที่เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 6 จำนวน 168 คน และ 183 คน ตามลำดับ ใช้โปรแกรม GENOVA ในการวิเคราะห์ข้อมูล ผลการเปรียบเทียบความตรงพบว่า ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ สูงกว่าความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับแบบประเมินผลสัมฤทธิ์ทางการเรียน ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ สูงกว่าความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบกับแบบประเมินผลสัมฤทธิ์ทางการเรียน ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับแบบประเมินผลสัมฤทธิ์ทางการเรียน ต่ำกว่าความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ แบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบและแบบประเมินผลสัมฤทธิ์ทางการเรียน ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบสอบวัดผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ แบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ และแบบประเมินผลสัมฤทธิ์ทางการเรียน ไม่มีความแตกต่างกัน ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบสอบวัดผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบกับแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบ มีความแตกต่างกันที่ระดับนัยสำคัญ 0.01 ความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบประเมินผลสัมฤทธิ์ทางการเรียนกับแบบสอบผลสัมฤทธิ์ทางการเรียนแบบเลือกตอบ ไม่มีความแตกต่างกัน และความตรงลู่เข้าระหว่างวิธีวัดด้วยแบบทดสอบผลสัมฤทธิ์ทางการเรียนแบบเติมคำตอบกับแบบประเมินผลสัมฤทธิ์ทางการเรียน ไม่มีความแตกต่างกัน

ปวีณา ปีอาทิตย์ (2545) ได้ทำการศึกษาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน โดยการวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) เปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนจากผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) และ 2) เพื่อ

หาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) กลุ่มตัวอย่างประกอบด้วย ผลงานการเขียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 จำนวน 117 ผลงาน และผู้ประเมินจำนวน 8 คน ที่ใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยก องค์ประกอบ การวิจัยนี้คำนวณค่าความเที่ยงจากค่าสัมประสิทธิ์การสรุปอ้างอิงด้วยโปรแกรมสำเร็จรูป GENOVA ผลการวิจัยพบว่า ภายใต้ชิ้นงานจำนวน 117 ชิ้น ผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.71484 และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.81410 2. ถ้าต้องการใช้ผู้ประเมิน 1 คน และกำหนดให้ค่าสัมประสิทธิ์การสรุปอ้างอิงของเกณฑ์การให้คะแนนแบบภาพรวมเป็น 0.2, 0.4, 0.6 และ 0.8 ใช้จำนวนชิ้นงานอย่างน้อย 4, 8, 16 และ 28 ชิ้น ตามลำดับ ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 และ 0.4 ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 ใช้จำนวนชิ้นงานอย่างน้อย 8 ชิ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.8 ใช้ชิ้นงานอย่างน้อย 12 ชิ้น ทั้งนี้ขึ้นอยู่กับผู้ปฏิบัติ ว่าจะเลือกใช้เกณฑ์การให้คะแนนให้เหมาะสมกับการเรียนการสอน

บุญประกอบ จายค่าย (2545) ได้ศึกษาการสร้างแบบวัดภาคปฏิบัติสำหรับโปรแกรมประมวลคำไมโครซอฟท์เวิร์ด โดยมีวัตถุประสงค์เพื่อ 1) สร้างแบบวัดภาคปฏิบัติสำหรับโปรแกรมประมวลคำไมโครซอฟท์เวิร์ด 2) เพื่อสร้างเกณฑ์ในการให้คะแนนในการวัดภาคปฏิบัติ 3) เพื่อหาคุณภาพของแบบวัดภาคปฏิบัติที่ผู้วิจัยสร้างขึ้น กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ภาคเรียนที่ 1 ปีการศึกษา 2545 ของโรงเรียนสาธิตมหาวิทยาลัยสงขลานครินทร์ จำนวน 60 คน แบ่งออกเป็น 2 กลุ่ม กลุ่มละ 30 คน โดยกลุ่มตัวอย่างดังกล่าวเลือกมาโดยวิธีสุ่มอย่างง่าย (simple random sampling) กลุ่มตัวอย่างกลุ่มที่ 1 ใช้ทดสอบเพื่อหาเวลาที่นักเรียนใช้ในการทำแบบวัด หาความถูกต้องชัดเจนของใบงาน ส่วนกลุ่มที่ 2 ใช้ทดสอบเพื่อหาค่าความเชื่อมั่นของแบบวัดและความเชื่อมั่นระหว่างผู้ประเมิน 2 คน เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือแบบวัดภาคปฏิบัติสำหรับโปรแกรมประมวลคำไมโครซอฟท์เวิร์ด ซึ่งประกอบด้วยใบงานสำหรับให้นักเรียนปฏิบัติแบบประเมินและเกณฑ์การให้คะแนน โดยแบบวัดดังกล่าวได้ผ่านการประเมินคุณภาพขั้นต้นโดยผู้เชี่ยวชาญ จำนวน 5 คน ซึ่งมีค่าดัชนีความสอดคล้องระหว่างข้อปฏิบัติกับจุดประสงค์เชิงพฤติกรรม ตั้งแต่ 0.80 - 1.00 การวิจัยครั้งนี้ได้แบบวัดภาคปฏิบัติดังนี้ 1. แบบวัดที่สร้างขึ้นในการวิจัยครั้งนี้ คือ แบบวัดภาคปฏิบัติสำหรับโปรแกรมประมวลคำไมโครซอฟท์เวิร์ด มีจำนวน 3 ฉบับ

ซึ่งประกอบด้วย แบบวัดการจัดการข้อมูล แบบวัดการสร้างตาราง และแบบวัดการทำจดหมายเวียน ซองจดหมายและป้ายผนึก แบบวัดการจัดการข้อมูลมี 2 ฉบับย่อยคือ แบบวัดการจัดการข้อมูล 1 และแบบวัดการจัดการข้อมูล 2 แบบวัดการทำจดหมายเวียน ซองจดหมายและป้ายผนึก มี 3 ฉบับย่อย คือแบบวัดการทำจดหมายเวียน แบบวัดการทำซองจดหมาย และแบบวัดการทำป้ายผนึก 2. ผลจากการสร้างเกณฑ์ในการให้คะแนนในการวัดภาคปฏิบัติ ผู้วิจัยได้สร้างเกณฑ์การให้คะแนนของแบบวัดแต่ละข้อเป็น 4 ระดับ คือ 0, 1, 2 และ 3 คะแนน ซึ่งการสร้างเกณฑ์การให้คะแนนได้ผ่านการพิจารณาของผู้เชี่ยวชาญ โดยมีค่าดัชนีความสอดคล้องระหว่างข้อปฏิบัติกับเกณฑ์การให้คะแนน ตั้งแต่ 0.80 - 1.00 และมีการกำหนดเกณฑ์การผ่านแบบวัดแต่ละฉบับโดยผู้เชี่ยวชาญ ได้เกณฑ์การผ่านแบบวัดที่ระดับร้อยละ 80 3. การหาคุณภาพของแบบวัดภาคปฏิบัติ โดยการหาค่าความเชื่อมั่นของแบบวัด ซึ่งคำนวณโดยสูตรสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) และหาค่าความเชื่อมั่นระหว่างผู้ประเมิน 2 คน คำนวณโดยใช้การประมาณค่าจากสูตรสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson Product Moment) ผลจากการหาคุณภาพของแบบวัดปรากฏดังนี้ 3.1 การหาค่าความเชื่อมั่นของแบบวัด ปรากฏดังนี้ แบบวัดการจัดการข้อมูล 1 มีค่าความเชื่อมั่นเท่ากับ 0.928 แบบวัดการจัดการข้อมูล 2 มีค่าความเชื่อมั่นเท่ากับ 0.987 แบบวัดการสร้างตาราง มีค่าความเชื่อมั่นเท่ากับ 0.987 แบบวัดการทำจดหมายเวียน มีค่าความเชื่อมั่นเท่ากับ 0.952 แบบวัดการทำซองจดหมาย มีค่าความเชื่อมั่นเท่ากับ 0.941 และแบบวัดการทำป้ายผนึก มีค่าความเชื่อมั่นเท่ากับ 0.931 3.2 การหาค่าความเชื่อมั่นระหว่างผู้ประเมิน 2 คน ปรากฏดังนี้ แบบวัดการจัดการข้อมูล 1 มีค่าความเชื่อมั่นเท่ากับ 0.928 แบบวัดการจัดการข้อมูล 2 มีค่าความเชื่อมั่นเท่ากับ 0.987 แบบวัดการสร้างตาราง มีค่าความเชื่อมั่นเท่ากับ 0.988 แบบวัดการทำจดหมายเวียน มีค่าความเชื่อมั่นเท่ากับ 0.952 แบบวัดการทำซองจดหมาย มีค่าความเชื่อมั่นเท่ากับ 0.937 และแบบวัดการทำป้ายผนึก มีค่าความเชื่อมั่นเท่ากับ 0.936

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

งานวิจัยที่เกี่ยวข้องกับทฤษฎีตอบสนองข้อสอบ

ธนวัฒน์ แสนสุข (2538) ได้ศึกษาการใช้จีอาร์เอ็ม จีพีซีเอ็ม และโมเดลโลจิสติกในการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบวัดที่มีวิธีการให้คะแนนแตกต่างกัน โดยในการวิจัยมีจุดมุ่งหมายเพื่อศึกษาการใช้ GRM, GPCM และโมเดลโลจิสติกในการเปรียบเทียบค่าค่าฟังก์ชันสารสนเทศของแบบวัดที่มีวิธีการให้คะแนนแบบทวิวิภาค (0,1) และแบบพหุวิภาค (1,2,3,4) ข้อมูลจากการตรวจให้คะแนนแบบทวิวิภาควิเคราะห์ตามโมเดลโลจิสติก 1,2 และ 3 พารามิเตอร์ ส่วนการตรวจให้คะแนนแบบพหุวิภาควิเคราะห์ตาม GRM และGPCM การวิจัยนี้ได้ศึกษากับข้อมูล 2 ชุด 1) ข้อมูลผลการวัดคุณลักษณะทั่วไปของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 6,300 คน จากสำนักงานทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ 2) ข้อมูลผลสัมฤทธิ์ทางคณิตศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 470 คน ที่ผู้วิจัยเก็บเอง ทำการตรวจสอบความสมบูรณ์ของข้อมูลโดยการวิเคราะห์ความเหมาะสมของผู้ตอบและข้อกระทงตาม RSM โดยโปรแกรม BIGSTEPS และวิเคราะห์เพื่อหาค่าฟังก์ชันสารสนเทศของแบบวัดโดยใช้โปรแกรม MULTILIG และ PARSCALE จากนั้นตรวจสอบความตรงเชิงโครงสร้างของแบบวัดคุณลักษณะทั่วไปโดยวิธีวิเคราะห์องค์ประกอบเชิงยืนยัน ด้วยโปรแกรม LISREL ผลการวิวัยพบว่า ผลการวิจัยพบว่าการให้คะแนนแบบพหุวิภาคเมื่อวิเคราะห์ตาม GRM ให้ค่าฟังก์ชันสารสรเทศสูงกว่าการตรวจให้คะแนนแบบทวิวิภาค สำหรับการตรวจให้คะแนนแบบพหุวิภาค วิเคราะห์ตาม GPCM และแบบทวิวิภาคยังไม่สามารถสรุปได้ว่าวิธีใดให้ค่าฟังก์ชันสารสนเทศสูงกว่ากัน

อรอนงค์ วิสาสะ (2540) ได้ศึกษาการเปรียบเทียบคุณภาพของแบบวัดฉบับสั้นที่ใช้ในการคัดเลือกข้อกระทง 3 วิธี โดยในการวิจัยมีวัตถุประสงค์เพื่อเปรียบเทียบคุณภาพแบบวัดฉบับสั้นที่ใช้วิธีการคัดเลือกข้อกระทง 3 วิธี มีวัตถุประสงค์เฉพาะเพื่อเปรียบเทียบความตรงเชิงโครงสร้างและความเที่ยงของแบบวัดฉบับสั้นที่ได้จากการคัดเลือกข้อกระทง 3 วิธี เครื่องมือที่ใช้ในการวิจัยคือแบบวัดนิสัยและทัศนคติในการเรียนชุดต้นฉบับของบราวน์ และ โฮลท์แมน (Brown and Holtzman) ซึ่งได้แปลเป็นฉบับภาษาไทยโดย ขจรสุดา หลักเพชร และแบบวัดฉบับสั้นทั้ง 3 ฉบับที่ได้จากการคัดเลือกข้อกระทงทั้ง 3 วิธี กลุ่มตัวอย่างคือนักเรียนชั้นมัธยมศึกษาปีที่ 2 ในเขตการศึกษาที่ 12 จังหวัดระยอง จำนวน 979 คน วิเคราะห์ข้อกระทงจากแบบวัดต้นฉบับเพื่อคัดเลือกข้อกระทงเป็นแบบวัดฉบับสั้น วิเคราะห์ค่าสถิติพื้นฐาน วิเคราะห์ความตรงเชิงโครงสร้างด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรมลิสเรล และวิเคราะห์ค่าความเที่ยงจากสูตรสัมประสิทธิ์แอลฟาของครอนบาค และค่าฟังก์ชันสารสนเทศของแบบวัดฉบับสั้นทั้ง 3 ฉบับ ผลการวิจัยพบว่า ความตรงเชิงโครงสร้างของแบบวัดทั้ง 3 ฉบับ มีความตรงเชิงโครงสร้างไม่

แตกต่างกัน ในส่วนของค่าความเที่ยงพบว่า ค่าความเที่ยงของแบบวัดฉบับสั้นโดยการคัดเลือกข้อกระทงจากดัชนีอำนาจจำแนกรายข้อตามไออาร์ที มีค่าความเที่ยงสูงสุด รองลงมาคือ แบบวัดฉบับสั้นที่ได้จากการคัดเลือกข้อกระทงจากน้ำหนักองค์ประกอบ และแบบวัดฉบับสั้นที่ได้จากการคัดเลือกข้อกระทงจากดัชนี รายข้อตามซีทีที แบบวัดฉบับสั้นทั้ง 3 ฉบับให้ค่าสารสนเทศสูงสำหรับผู้ตอบที่มีความสามารถระดับปานกลาง (θ ระหว่าง -2 ถึง 2) โดยแบบวัดฉบับสั้นที่คัดเลือกข้อกระทงจากน้ำหนักองค์ประกอบให้ค่าสารสนเทศของแบบวัดสูงสุด รองลงมาเป็นแบบวัดฉบับสั้นที่คัดเลือกข้อกระทงจากดัชนีอำนาจจำแนกรายข้อตามซีทีที และแบบวัดฉบับสั้นที่คัดเลือกข้อกระทงจากดัชนีอำนาจจำแนกรายข้อตามไออาร์ที

พรทิพย์ เพชรกาศ (2543) ได้ศึกษาการเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบเลือกตอบแบบถูกผิด ที่ตรวจให้คะแนนต่างวิธี โดยการวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบเลือกตอบแบบถูกผิดที่ตรวจให้คะแนนต่างกัน 3 วิธีคือ วิธี 0-1 , วิธี Pass-3 และวิธี MTF โดยใช้ทฤษฎีการตอบสนองข้อสอบ (IRT) โมเดลโลจีสติค 3 พารามิเตอร์ กลุ่มตัวอย่างที่ใช้ในการวิจัยนี้เป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2542 ของโรงเรียนในสังกัดสำนักงานการประถมศึกษาจังหวัดพัทลุง จำนวน 1,013 คน ได้มาโดยการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการรวบรวมข้อมูลเป็นแบบทดสอบเลือกตอบแบบถูกผิดที่มี 4 ตัวเลือก วิชาสร้างเสริมประสบการณ์ชีวิต จำนวน 20 ข้อ ที่ผู้วิจัยสร้างขึ้น ผลการวิจัยพบว่า : 1. ค่าฟังก์ชันสารสนเทศของแบบทดสอบเลือกตอบแบบถูกผิดที่ตรวจให้คะแนน 3 วิธี มีค่าแตกต่างกันทุกระดับความสามารถที่เปรียบเทียบ 2. ที่ระดับความสามารถต่ำมากถึงปานกลาง (-4.00 0.70) แบบทดสอบที่ตรวจให้คะแนนวิธี MTF มีค่าฟังก์ชันสารสนเทศสูงกว่าแบบทดสอบที่ตรวจให้คะแนนวิธี Pass-3 และวิธี 0-1 3. ที่ระดับความสามารถปานกลางค่อนข้างสูงถึงสูงมาก (0.70 4.00) แบบทดสอบที่ตรวจให้คะแนนวิธี 0-1 มีค่าฟังก์ชันสารสนเทศสูงกว่าแบบทดสอบที่ตรวจให้คะแนนวิธี MTF และวิธี Pass-3

สุพจน์ เกิดสุวรรณ์ (2545) พัฒนาวิธีการวัดความรู้บางส่วนของผู้ตอบแบบสอบเลือกตอบ การดำเนินการวิจัยประกอบด้วย 2 ขั้นตอน ขั้นตอนแรกเป็นการพัฒนาวิธีการตอบและตรวจให้คะแนน โดยประยุกต์วิธีของคูมบ์ ประยุกต์วิธีของอาร์โนลด์ และประยุกต์วิธีของเดรสเซลและชมิด ขั้นตอนที่สอง วิเคราะห์คุณภาพวิธีการตอบและตรวจให้คะแนนที่พัฒนาขึ้น เพื่อเปรียบเทียบคุณภาพกับวิธีการตอบและตรวจให้คะแนนตามวิธีของคูมบ์ วิธีของอาร์โนลด์ วิธีของเดรสเซลและชมิด และวิธีการตรวจให้คะแนนแบบประเพณีนิยม ในด้านความตรงตามเกณฑ์ ความตรงเชิง

โครงสร้าง ความเที่ยงแบบความสอดคล้องภายใน ค่าฟังก์ชันสารสนเทศเฉลี่ยของแบบสอบ และค่าประสิทธิภาพสัมพัทธ์เฉลี่ยของแบบสอบ เครื่องมือที่ใช้ในการวิจัยเป็นแบบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน ระดับมัธยมศึกษาตอนต้น โดยการสุ่มกลุ่มตัวอย่างนักเรียนระดับมัธยมศึกษาปีที่ 3 จากโรงเรียนมัธยมศึกษาในจังหวัดสระบุรี ปีการศึกษา 2544 จำนวน 15 โรงเรียน รวมทั้งสิ้น 946 คน ข้อมูลที่ได้นำมาตรวจสอบความตรงเชิงโครงสร้าง โดยการวิเคราะห์เมตริกพหุลักษณะ-พหุวิธี และวิเคราะห์ตัวประกอบเชิงยืนยันด้วยโปรแกรม LISREL ตรวจสอบความตรงตามเกณฑ์สัมพันธ์จากการวิเคราะห์ สหสัมพันธ์ของคะแนนกับเกณฑ์ วิเคราะห์ค่าความเที่ยงแบบความสอดคล้องภายใน และวิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบสอบ ด้วยโปรแกรม BILOG และ PARSCALE ผลการวิจัยสรุปได้ดังนี้ 1. วิธีที่ประยุกต์จากวิธีของคูมบ์ ให้ค่าฟังก์ชันสารสนเทศเฉลี่ยของแบบสอบ และประสิทธิภาพสัมพัทธ์เฉลี่ยของแบบสอบสูงกว่าวิธีอื่นๆ รองลงมาคือวิธีของอาร์โนลด์ และวิธีที่ประยุกต์จากวิธีของเดรสเซลและชมิด ตามลำดับ ส่วนวิธีประเพณีนิยม จะให้ค่าฟังก์ชันสารสนเทศเฉลี่ยของแบบแบบสอบ และประสิทธิภาพสัมพัทธ์เฉลี่ยของแบบสอบต่ำกว่าทุกวิธี 2.วิธีของอาร์โนลด์มีความตรงเชิงโครงสร้าง สูงกว่าวิธีอื่นๆ รองลงมาคือ วิธีที่ประยุกต์จากวิธีของคูมบ์และวิธีของคูมบ์ ตามลำดับ ส่วนวิธีประเพณีนิยมมีความตรงเชิงโครงสร้างต่ำกว่าทุกวิธี 3 วิธีของคูมบ์ และวิธีที่ประยุกต์จากวิธีของคูมบ์มีค่าความตรงตามเกณฑ์สัมพันธ์ สูงกว่าวิธีที่ประยุกต์จากวิธีของอาร์โนลด์ 4. วิธีที่ประยุกต์จากวิธีของอาร์โนลด์ มีค่าความเที่ยงสูงกว่าวิธีอื่นๆ รองลงมาคือวิธีของอาร์โนลด์ และวิธีที่ประยุกต์จากวิธีของเดรสเซลและชมิด ตามลำดับ 5.คุณภาพของวิธีการตอบและการตรวจให้คะแนน เมื่อพิจารณาตามเกณฑ์โดยรวมทุกด้านพบว่า วิธีที่ประยุกต์จากวิธีของคูมบ์ มีคุณภาพดีกว่าวิธีอื่นๆ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาและพัฒนาเกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ด้วยวิธีของโนแวคและโกวิน และวิธีของบอลท์ และเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ เมื่อตรวจโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และใช้วิธีการให้คะแนนของบอลท์ โดยพิจารณาจากความตรงตามสภาพ ความตรงเชิงโครงสร้าง ความเที่ยงซึ่งแทนด้วยสัมประสิทธิ์สรุปอ้างอิง ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ และวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสม ซึ่งผู้วิจัยได้เสนอลำดับขั้นตอนการวิจัยดังนี้

ตอนที่ 1 ประชากรและกลุ่มตัวอย่าง

ตอนที่ 2 เครื่องมือที่ใช้ในการวิจัย

ตอนที่ 3 การสร้างเครื่องมือที่ใช้ในการวิจัย

ตอนที่ 4 การนำเครื่องมือที่ใช้ในการวิจัยไปเก็บรวบรวมข้อมูล

ตอนที่ 5 การวิเคราะห์ข้อมูล

ตอนที่ 1 ประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นผู้ประเมินซึ่งเป็นครูสอนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1-3 และนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ

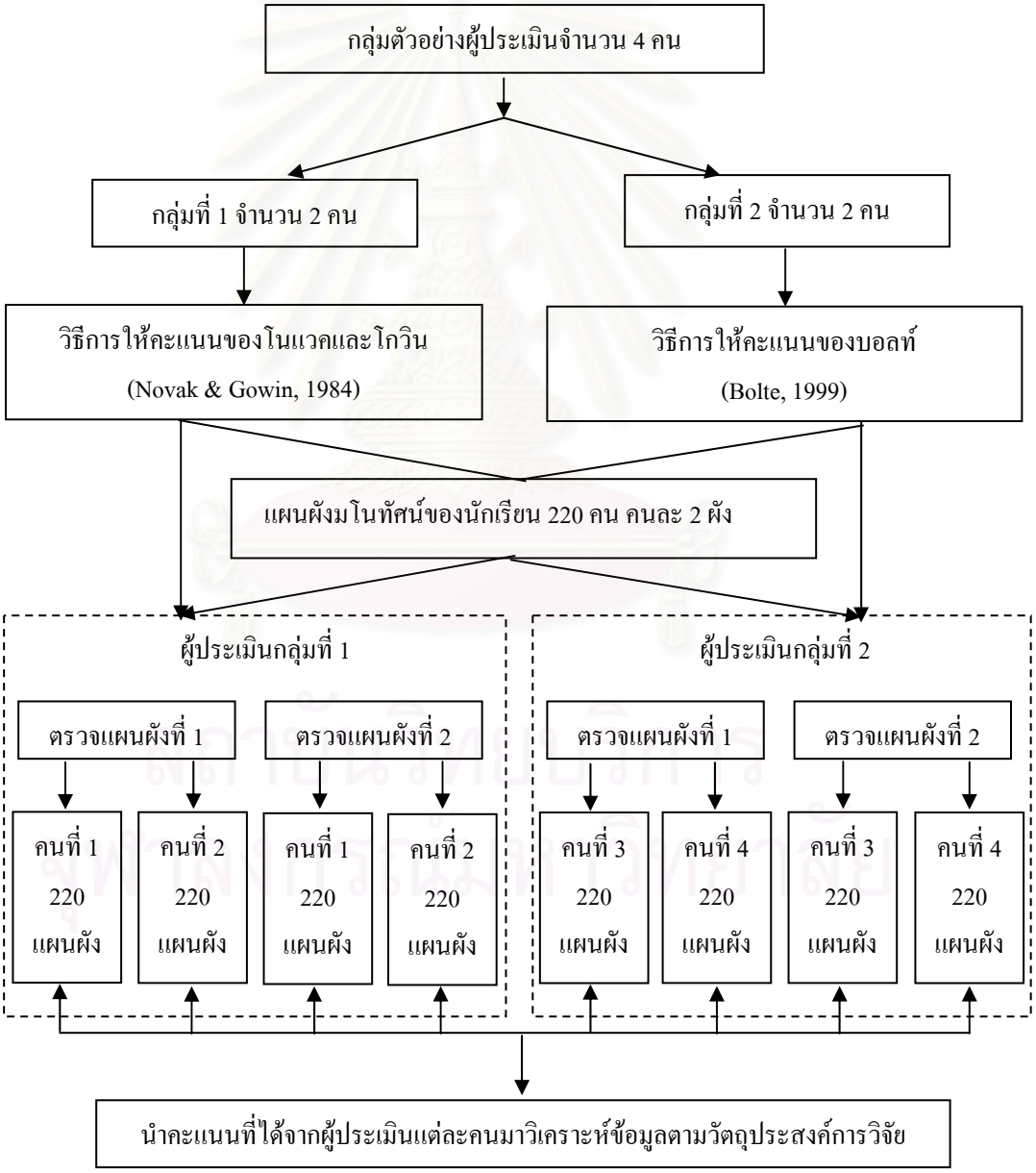
กลุ่มตัวอย่าง

ในการวิจัยครั้งนี้มีกลุ่มตัวอย่างที่ศึกษาประกอบด้วย

3. กลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 โรงเรียนสุรนารีวิทยา สำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต1 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ จำนวน 220 คน

4. กลุ่มตัวอย่างผู้ประเมิน เนื่องจากในการวิจัยครั้งนี้ต้องใช้ครูผู้สอนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1-3 เป็นผู้ประเมินแผนผังมโนทัศน์ซึ่งต้องใช้เวลาและความสนใจ ตั้งใจในการประเมิน ผู้วิจัยทำการเลือกครูผู้ประเมินโดยวิธีการเลือกแบบเจาะจง (purposive sampling) เพราะ

ต้องหาผู้ประเมินที่เต็มใจและมีเวลาช่วยประเมินอย่างเต็มที่ มีประสบการณ์ในการสอนอย่างน้อย 3 ปี ครูที่ได้รับเลือกเป็นกลุ่มตัวอย่าง รวมทั้งหมด 4 คน แบ่งครู 4 คน เป็น 2 กลุ่ม กลุ่มละ 2 คน โดยวิธีสุ่มอย่างง่ายโดยการจับสลาก ซึ่งผู้ประเมินกลุ่มที่ 1 จะใช้วิธีการให้คะแนนของโนแวคและโกวิน (Novak & Gowin, 1984) โดยผู้ประเมินทุกคนตรวจให้คะแนนแผนผังที่ 1 และแผนผังที่ 2 และผู้ประเมินกลุ่มที่ 2 จะใช้วิธีการให้คะแนนของบอลท์ (Bolte, 1999) โดยผู้ประเมินทุกคนตรวจให้คะแนนแผนผังที่ 1 และแผนผังที่ 2 ดังนั้น ผู้ประเมินแต่ละคนจะตรวจผลงานของนักเรียนจำนวน 220 คน จำนวน 440 ชิ้น โดยผู้วิจัยขอสรุปขั้นตอน ดังแผนภาพที่ 3.1



แผนภาพที่ 3.1 แสดงการตรวจผลงานนักเรียนของกลุ่มตัวอย่างผู้ประเมิน

ตอนที่ 2 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการดำเนินการวิจัยประกอบด้วย

1. แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) เป็นแบบประเมินที่กำหนดคำมโนทัศน์มาให้ แล้วให้นักเรียนสร้างแผนผังลงในกระดาษที่แจกให้ โดยในการสร้างแผนผังนักเรียนสามารถเพิ่มคำมโนทัศน์ได้ตามความคิดของตนเอง พร้อมทั้งลากเส้นเชื่อมโยงและเขียนคำเชื่อมแสดงความสัมพันธ์ระหว่างมโนทัศน์ให้สมบูรณ์มีจำนวน 2 แผนผัง แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) ดังกล่าวเป็นแบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์ มีเนื้อหาครอบคลุม 2 เรื่อง คือ สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1

2. แบบทดสอบประเภทเลือกตอบ เป็นแบบวัดความคิดรวบยอดวิชาคณิตศาสตร์มีเนื้อหาครอบคลุม 2 เรื่อง คือ สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1

ตอนที่ 3 การสร้างเครื่องมือที่ใช้ในการวิจัย

ในการสร้างเครื่องมือที่ใช้ในการวิจัยนี้ ผู้วิจัยขอนำเสนอเป็น 5ระยะดังนี้

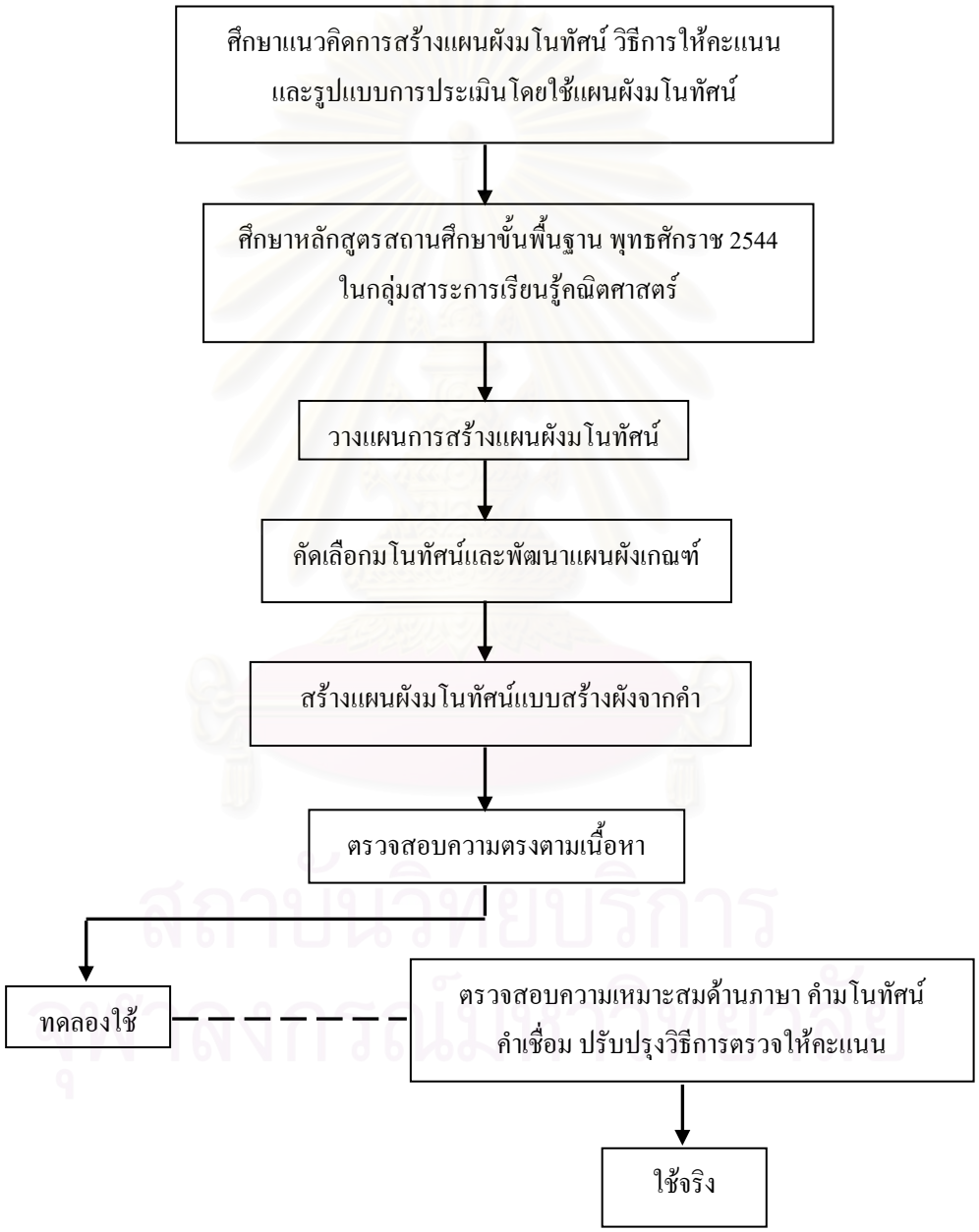
ตอนที่ 3.1 การสร้างแบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map)

ตอนที่ 3.2 การสร้างแบบทดสอบประเภทเลือกตอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 3.1 การสร้างแบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map)

ในการสร้างแบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ผู้วิจัยได้ทำตามขั้นตอนในแผนภาพที่ 3.2 ซึ่งนำเสนอในรายละเอียดดังนี้



แผนภาพที่ 3.2 การสร้างแบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์

1. ศึกษาแนวคิดการสร้างแผนผังมโนทัศน์ รูปแบบการประเมินโดยใช้แผนผังมโนทัศน์ ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการสร้างแผนผังมโนทัศน์ โดยศึกษาวิธีการสร้างและองค์ประกอบของแผนผังมโนทัศน์มโนทัศน์ ซึ่งสามารถสรุปองค์ประกอบของแผนผังมโนทัศน์ที่นำมาใช้ในการศึกษาวิจัยได้ 3 องค์ประกอบ คือ คำมโนทัศน์ คำเชื่อม และเส้นเชื่อมระหว่างมโนทัศน์

2. ศึกษาวิธีการให้คะแนนแผนผังมโนทัศน์

ผู้วิจัยศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการสร้างแผนผังมโนทัศน์ โดยศึกษาวิธีการให้คะแนนของโนแวคและโกวิน (Novak & Gowin, 1984) และวิธีการให้คะแนนของบอลท์ (Bolte, 1999)

3. ศึกษาหลักสูตรสถานศึกษาขั้นพื้นฐาน พุทธศักราช 2544

ผู้วิจัยกับครูผู้สอนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1 ร่วมกันศึกษาและวิเคราะห์หลักสูตรสถานศึกษาขั้นพื้นฐาน พุทธศักราช 2544 ในกลุ่มสาระการเรียนรู้คณิตศาสตร์ช่วงชั้นที่ 1 (ม.1-ม.3) เพื่อคัดเลือกเนื้อหาคณิตศาสตร์ของระดับชั้นมัธยมศึกษาปีที่ 1 มาสร้างแบบประเมิน โดยเลือกเนื้อหาวิชาคณิตศาสตร์ เรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1

4. วางแผนการสร้างแบบประเมิน

การสร้างแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์วิชาคณิตศาสตร์นี้ ผู้วิจัยได้ประยุกต์แนวคิดในการสร้างแผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) ของรุซ-พรีโมและคณะ (Ruiz-Primo et al, 2001) สาเหตุที่ผู้วิจัยได้นำแผนผังมโนทัศน์รูปแบบนี้มาใช้เนื่องจาการการศึกษางานวิจัยดังกล่าว พบว่าแผนผังมโนทัศน์รูปแบบนี้สามารถนำไปใช้สะท้อนความคิดรวบยอด และความเข้าใจเนื้อหา ในบทเรียนของนักเรียนได้ดีกว่าแบบอื่น (Yin et al, 2004)

5. คัดเลือกมโนทัศน์และพัฒนาแผนผังเกณฑ์

การสร้างแผนผังมโนทัศน์เกณฑ์ (criterion map) เป็นขั้นตอนที่สำคัญมาก เพราะแผนผังเกณฑ์เป็นแผนผังหลักหรือแผนผังเฉลยเพื่อใช้เป็นแนวทางในการสร้างแผนผังมโนทัศน์ในการสร้างแผนผังมโนทัศน์เกณฑ์ ผู้วิจัยร่วมกับครูผู้สอนวิชาคณิตศาสตร์ทำการวิเคราะห์เนื้อหา

ในเรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1 ซึ่งเป็นเนื้อหาที่ใช้ในการสร้างแผนผังมโนทัศน์ จากหลักสูตรสถานศึกษาขั้นพื้นฐาน พุทธศักราช 2544 หนังสือเรียนคณิตศาสตร์ คู่มือครู และตำราอ่านประกอบอื่นๆ โดยแยกเป็นเนื้อหาหลักและเนื้อหาย่อย พร้อมทั้งกำหนดจำนวนแผนผังมโนทัศน์ที่จะสร้างเป็นแบบประเมิน ให้ครอบคลุมเนื้อหาที่ต้องการวัด เมื่อได้เนื้อหาสาระแล้วผู้วิจัยและครูผู้สอนได้ร่วมกันสร้างแผนผังเกณฑ์ตามมโนทัศน์ที่มุ่งวัดโดยมีขั้นตอนดังนี้

5.1 คัดเลือกมโนทัศน์ ผู้วิจัยและครูผู้สอนร่วมกันสำรวจและรวบรวมมโนทัศน์หลัก มโนทัศน์รอง หรือมโนทัศน์ย่อยของแต่ละเนื้อหาในเรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1 จากคู่มือครู หนังสือเรียน และหนังสือคณิตศาสตร์ทั่วๆไป โดยมโนทัศน์ที่คัดเลือกจะต้องคำนึงถึงเนื้อหาที่ใช้สอนสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 แล้วคัดลอกคำมโนทัศน์เหล่านั้นลงในกระดาษที่เตรียมไว้ โดยแยกคำมโนทัศน์เป็นหมวดหมู่ตามเนื้อหาที่มุ่งวัด

5.2 จัดลำดับมโนทัศน์ นำคำมโนทัศน์หลัก มโนทัศน์รองของเนื้อหาแต่ละเรื่องที่เขียนไว้ในกระดาษข้อ 5.1 มาสร้างเป็นแผนผังเกณฑ์ของแบบประเมินแต่ละชุด การจัดลำดับมโนทัศน์ทำได้โดยการเลือกคำมโนทัศน์หลักของแต่ละเรื่องวางไว้บนสุด แล้วจัดคำมโนทัศน์รองมโนทัศฯย่อยเรียงลดหลั่นกันไป โดยใช้เกณฑ์การเลือกคำมโนทัศน์กลุ่มเดียวกันอยู่ระดับเดียวกัน

5.3 เชื่อมโยงความสัมพันธ์ระหว่างมโนทัศน์ เมื่อจัดคำมโนทัศน์เสร็จแล้วผู้วิจัยลากลูกศรเชื่อมโยงระหว่างคำมโนทัศน์ที่จัดเรียงไว้ ทั้งนี้เพื่อแสดงถึงทิศทางความสัมพันธ์ พร้อมทั้งเขียนคำเชื่อมกำกับไว้ที่เส้นแต่ละเส้นเพื่อแสดงถึงความสัมพันธ์ระหว่างมโนทัศน์

5.4 ผู้วิจัยตรวจสอบความถูกต้องของคำมโนทัศน์ คำเชื่อม เส้นเชื่อมโยงอีกครั้งโดยผ่านการเห็นชอบจากครูผู้สอนได้ช่วยพิจารณาความเหมาะสมด้วย แล้วนำมาจัดพิมพ์เพื่อใช้เป็นแผนผังเกณฑ์ซึ่งเป็นแผนผังหลักที่จะใช้สร้างเป็นแผนผังมโนทัศน์ต่อไป

6. สร้างแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์

การสร้างแบบประเมินความคิดรวบยอดโดยใช้แผนผังมโนทัศน์ ผู้วิจัยนำแผนผังมโนทัศน์เกณฑ์ที่สร้างไว้ในข้อ 5 มาใช้เป็นแผนผังหลักในการสร้างแบบประเมิน ซึ่งเนื้อหาหรือมโนทัศน์ที่มุ่งวัดทั้งหมดจะนำมาสร้างแบบประเมินประเภทแผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) ซึ่งเป็นแบบประเมินที่นำคำมโนทัศน์ทั้งหมดในแผนผังเกณฑ์มาใช้เป็นคำมโนทัศน์ที่กำหนดให้ เพื่อให้นักเรียนนำคำมโนทัศน์เหล่านั้นมาสร้างเป็นแผนผังมโนทัศน์ ซึ่งในการสร้างแผนผังมโนทัศน์จะให้อิสระแก่นักเรียนเต็มที่ โดยผู้วิจัยไม่ได้กำหนดรูปแบบของ

แผนผังมาให้ นักเรียนสามารถสร้างสรรค์รูปแบบของแผนผังได้ตามต้องการ นอกจากนี้นักเรียนยังสามารถเพิ่มคำมโนทัศน์ที่นอกเหนือจากคำมโนทัศน์ที่กำหนดให้ได้ ซึ่งในการสร้างแผนผังมโนทัศน์นักเรียนต้องลากลูกศรเพื่อแสดงทิศทางของคำมโนทัศน์ โดยแผนผังมโนทัศน์ที่สร้างขึ้นจะประกอบด้วย 3 องค์ประกอบ คือ คำมโนทัศน์ คำเชื่อม และการลากเส้นโยงความสัมพันธ์ระหว่างมโนทัศน์ (แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ แสดงไว้ที่ภาคผนวก ช)

7. การตรวจสอบความตรงตามเนื้อหาโดยผู้ทรงคุณวุฒิ

การตรวจสอบความตรงตามเนื้อหาโดยพิจารณาจากความสอดคล้องระหว่างเนื้อหาที่มุ่งวัดกับคำมโนทัศน์ (item objective congruence: IOC) ผู้วิจัยนำแผนผังมโนทัศน์เกณฑ์แบบทดสอบชนิดเลือกตอบ ที่ได้ผ่านการตรวจสอบโดยอาจารย์ที่ปรึกษาวิทยานิพนธ์แล้วไปให้ผู้ทรงคุณวุฒิทางการสอนวิชาคณิตศาสตร์ 3 ท่านตรวจสอบความครอบคลุมของโครงสร้างเนื้อหาและตรวจสอบความสอดคล้องระหว่างเนื้อหาที่มุ่งวัดกับคำมโนทัศน์ที่ต้องการวัดในแผนผังเกณฑ์

ผลการตรวจสอบความครอบคลุมของโครงสร้างเนื้อหา และตรวจสอบความสอดคล้องระหว่างเนื้อหาที่มุ่งวัดกับคำมโนทัศน์ที่ต้องการวัดในแผนผังเกณฑ์ ผู้ทรงคุณวุฒิได้ให้คำแนะนำในการปรับปรุงแผนผังมโนทัศน์ที่พัฒนาขึ้นโดยมีรายละเอียดดังนี้

ตารางที่ 3.1 ผลการตรวจสอบความตรงตามเนื้อหาโดยวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) ระหว่างเนื้อหาที่มุ่งวัดกับคำมโนทัศน์ที่ต้องการวัดในแผนผังเกณฑ์ของผู้ทรงคุณวุฒิ 3 ท่าน

รายการประเมิน	ดัชนีความสอดคล้องของความคิดเห็นของผู้ทรงคุณวุฒิ	
	สมบัติของจำนวนนับ	ระบบจำนวนเต็ม
1. ความสอดคล้องระหว่างคำ มโนทัศน์กับเนื้อหาที่มุ่งวัด	1.00	1.00
2. ความถูกต้องของคำเชื่อมระหว่างมโนทัศน์	0.66	0.66
3. ความถูกต้องของเส้นเชื่อมระหว่างมโนทัศน์	0.66	0.66
4. ความเหมาะสมของด้านการใช้ภาษา และคำชี้แจงต่างๆ	1.00	1.00

จากตารางที่ 3.1 พบว่าผู้ทรงคุณวุฒิมีความเห็นสอดคล้องกัน โดยเนื้อหาที่มุ่งวัดกับคำมโนทัศน์ที่ต้องงการวัดในแผนผังเกณฑ์และแผนผังมโนทัศน์จำนวน 2 ชุด สามารถวัดได้ตรงตามเนื้อหาและมีค่าดัชนีความสอดคล้อง (IOC) ตั้งแต่ 0.66-1.00 ซึ่งถือว่าผู้ทรงคุณวุฒิส่วนใหญ่มีความเห็นสอดคล้องกันมากกว่าร้อยละ 50 นอกจากนี้ผู้ทรงคุณวุฒิได้พิจารณาถึงรูปแบบของแผนผังมโนทัศน์ การใช้ภาษา และคำชี้แจงต่างๆ โดยมีรายละเอียดของความเห็นและผลการพิจารณาดังนี้

ด้านรูปแบบของแผนผังมโนทัศน์ ผู้ทรงคุณวุฒิได้แนะนำว่าแผนผังมโนทัศน์บางเรื่องมีเนื้อหามากเกินไป และมีเส้นเชื่อมโยงระหว่างมโนทัศน์มาก ควรปรับเส้นเชื่อมโยงที่ลากจากมโนทัศน์หลักไปยังมโนทัศน์อื่นๆ ให้ดูง่ายขึ้น ซึ่งในกรณีที่ใช้คำเชื่อมเหมือนกันหรือมโนทัศน์ระดับเดียวกัน ควรลากเส้น และเขียนคำเชื่อมกำกับไว้เพียงเส้นเดียว เพื่อลดจำนวนคำเชื่อมลง

สำหรับตัวอักษรที่ใช้ในแผนผังมโนทัศน์ ควรใช้ตัวอักษรที่มีขนาดใหญ่ขึ้น เพื่อให้นักสามารถมองเห็นได้ชัดเจนยิ่งขึ้น

ด้านคำมโนทัศน์ ควรใช้คำลักษณะเดียวกัน คือ ควรใช้คำนามให้หมดทุกคำ เพื่อเป็นลักษณะคำมโนทัศน์ในแนวเดียวกัน

ด้านคำเชื่อม ปรับคำเชื่อมบางคำให้เหมาะสม

ส่วนในด้านคำชี้แจงและคำสั่ง มีความชัดเจนและเข้าใจดี

8. การทดลองใช้เครื่องมือ

ผู้วิจัยได้นำแบบประเมินที่ได้ปรับปรุงตามคำแนะนำของผู้ทรงคุณวุฒิ ไปทดลองใช้กับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ที่ไม่ใช่กลุ่มตัวอย่างจำนวน 15 คน วัตถุประสงค์ของการทดลองใช้แบบประเมินครั้งนี้ เพื่อตรวจสอบความเหมาะสมด้านภาษา และความเข้าใจในคำชี้แจงของแผนผังมโนทัศน์

ผลการทดลองใช้เครื่องมือพบว่า นักเรียนส่วนใหญ่มีความเข้าใจสำนวน ถ้อยคำ ภาษา และคำชี้แจงที่ใช้ในแผนผังมโนทัศน์เป็นอย่างดี และพบว่าแผนผังมโนทัศน์แบบสร้างผังจากคำนั้น นักเรียนส่วนใหญ่เห็นว่าการสร้างแผนผังเป็นเรื่องยาก โดยเฉพาะการเลือกใช้คำเชื่อม จากการสังเกตพฤติกรรมนักเรียนพบว่าส่วนใหญ่ยังไม่สามารถจัดประเภทของคำ และไม่สามารถหาคำมาเชื่อมระหว่างมโนทัศนีได้ถูกต้อง นอกจากนี้ยังพบว่านักเรียนใช้คำมโนทัศน์ไม่ครบทุกคำ และนักเรียนไม่สามารถยกตัวอย่างเข้าไปในแผนผังมโนทัศน์ได้ จากปัญหาที่พบดังกล่าวจึงทำให้ผู้วิจัยมีความจำเป็นอย่างยิ่งที่ต้องเตรียมความพร้อมในการสร้างแผนผังมโนทัศน์ให้แก่นักเรียนมากขึ้น โดยเฉพาะอย่างยิ่งในการนำมาจัดกิจกรรมการเรียนการสอน เพราะหากนักเรียนไม่มีความชำนาญ

หรือไม่เข้าวิธีการสร้างแผนผังมโนทัศน์ ก็อาจส่งผลต่อความคลาดเคลื่อนในการวัดมโนทัศน์ของนักเรียนได้

ตอนที่ 3.2 การสร้างแบบทดสอบประเภทเลือกตอบ

แบบสอบที่สร้างขึ้นเป็นแบบทดสอบประเภทเลือกตอบ 4 ตัวเลือก ในสาระการเรียนรู้คณิตศาสตร์ เรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1 โดยมีวิธีการสร้างดังนี้

1. ศึกษาหลักสูตร สาระการเรียนรู้คณิตศาสตร์ มาตรฐานการเรียนรู้ช่วงชั้นและผลการเรียนรู้ที่คาดหวัง เรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม ในชั้นมัธยมศึกษาปีที่ 1 จาก แผนการจัดการเรียนรู้ และเอกสารที่เกี่ยวข้อง และจัดทำตารางแสดงแผนผังข้อสอบ (ตารางแสดงแผนผังข้อสอบ ได้แสดงไว้ในภาคผนวก ข)

2. สร้างเครื่องมือตามมาตรฐานการเรียนรู้และผลการเรียนรู้ที่คาดหวัง โดยแต่ละมาตรฐานการเรียนรู้จะสร้างข้อสอบให้มีจำนวนมากกว่าที่ต้องการ ซึ่งเครื่องมือที่สร้างขึ้นเป็นแบบทดสอบประเภทเลือกตอบ 4 ตัวเลือก จำนวน 50 ข้อ

3. นำแบบตรวจสอบคุณภาพขั้นต้น โดยให้ผู้เชี่ยวชาญทางคณิตศาสตร์ จำนวน 3 ท่าน พิจารณาความสอดคล้องของข้อสอบกับมาตรฐานการเรียนรู้ เพื่อตรวจสอบความตรงตามเนื้อหา โดยหาค่าดัชนีความสอดคล้องของข้อสอบ (Item Objective Congruence : IOC) ข้อสอบที่ใช้ได้มีค่า IOC ตั้งแต่ 0.66 ถึง 1.00 พร้อมทั้งปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ

4. นำแบบสอบที่ได้จากข้อ 3 ไปทดลองกับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ที่ไม่ใช่กลุ่มตัวอย่างจำนวน 55 คน จากโรงเรียนบุญวัฒนา จังหวัดนครราชสีมา เนื่องจากกลุ่มทดลองเครื่องมือดังกล่าวมีลักษณะคล้ายกลุ่มตัวอย่างสำหรับการวิจัยครั้งนี้

5. ตรวจให้คะแนนโดยให้คะแนนข้อที่ตอบถูก 1 คะแนน ข้อที่ตอบผิด 0 คะแนน

6. นำผลจากข้อ 5. มาวิเคาะห์คุณภาพข้อสอบ โดยใช้โปรแกรมวิเคราะห์ข้อสอบ TAP (Test Analysis Program) วิเคราะห์ข้อสอบรายข้อ คัดเลือกข้อสอบที่มีค่าความยาก (p) 0.20-0.80 ค่าอำนาจจำแนก (r) ตั้งแต่ 0.20 ขึ้นไป (ค่าความยากและค่าอำนาจจำแนก จากการวิเคราะห์ข้อสอบได้แสดงไว้ในภาคผนวก ค)

7. พิจารณาจำนวนข้อสอบให้เหมาะสมกับจุดประสงค์การเรียนรู้ เลือกข้อสอบสำหรับแบบทดสอบฉบับละ 13 ข้อ จำนวน 2 ฉบับ (สมบัติของจำนวนนับ และระบบจำนวนเต็ม) รวมทั้งสิ้น 26 ข้อ

ตอนที่ 4 การนำเครื่องมือที่ใช้ในการวิจัยไปเก็บรวบรวมข้อมูล

ในการเก็บรวบรวมข้อมูลครั้งนี้ แบ่งออกเป็น 2 ระยะ คือ

ตอนที่ 4.1 การเตรียมความพร้อมให้กับผู้ประเมิน

ตอนที่ 4.2 การเก็บรวบรวมข้อมูลเพื่อนำมาศึกษาคุณภาพด้านความเที่ยงเมื่อแทนค่าด้วยสัมประสิทธิ์การสรุปอ้างอิง คุณภาพด้านค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบและความตรงเชิงโครงสร้าง

ตอนที่ 4.3 การเก็บรวบรวมข้อมูลเพื่อนำมาศึกษาคุณภาพด้านความตรงตามสภาพ

ซึ่งมีขั้นตอนการเก็บรวบรวมข้อมูลดังนี้

ตอนที่ 4.1 การเตรียมความพร้อมให้กับผู้ประเมิน

1. ผู้วิจัยดำเนินการให้ความรู้เรื่องแผนผังมโนทัศน์ เพื่อสร้างความเข้าใจให้กับครูผู้สอนวิชาคณิตศาสตร์ทั้ง 4 ท่าน โดยหัวข้อการบรรยายประกอบด้วย การให้นิยามเชิงปฏิบัติการของแผนผังมโนทัศน์ องค์ประกอบของแผนผังมโนทัศน์ ประเภทของแผนผังมโนทัศน์และวิธีการสร้างแผนผังมโนทัศน์

2. ผู้วิจัยได้ชี้แจงเพิ่มเติมให้ผู้ประเมินกลุ่มแรกทราบรายละเอียดของการตรวจด้วยเกณฑ์การให้คะแนนของโนแวคและโกวิน และผู้ประเมินกลุ่มที่ 2 ทราบรายละเอียดเกณฑ์การให้คะแนนของบอลท์ โดยอธิบายเกณฑ์การให้คะแนนตามตัวบ่งชี้และระดับคุณภาพ ดังนี้

2.1 เกณฑ์การให้คะแนนของโนแวคและโกวิน จะพิจารณาแผนผังมโนทัศน์ในด้านคุณภาพ 4 ด้าน คือ พิจารณาจากประพจน์ (propositions) การจัดลำดับขั้นตอน (hierarchy) การเชื่อมระหว่างมโนทัศน์ (cross link) และตัวอย่าง (examples) ซึ่งแต่ละด้านจะแบ่งคะแนนออกเป็น 5 ระดับ จาก 0-4 ดังนั้นแผนผังมโนทัศน์ที่ตรวจให้ด้วยเกณฑ์การให้คะแนนของโนแวคและโกวินจะมีคะแนนเต็ม 16 คะแนน

2.2 เกณฑ์การให้คะแนนของบอลท์ จะพิจารณาให้คะแนนในด้านการจัดแผนผังมโนทัศน์ (organization) 6 คะแนน แบ่งคะแนนออกเป็น 7 ระดับ จาก 0-6 และความถูกต้อง (accuracy) 4 คะแนน แบ่งคะแนนออกเป็น 5 ระดับ จาก 0-4 ดังนั้นแผนผังมโนทัศน์ที่ตรวจด้วยเกณฑ์การให้คะแนนของบอลท์จะมีคะแนนเต็ม 10 คะแนน

3. ชี้แจงการประเมินแผนผังมโนทัศน์ว่า ผู้ประเมินต้องตรวจให้คะแนนแผนผังมโนทัศน์ของนักเรียน 220 คน คนละ 2 แผนผัง รวม 440 แผนผัง โดยผู้ประเมินตรวจแผนผังที่ 1 จำนวน 220 คนก่อน แล้วจึงตรวจแผนผังที่ 2 จำนวน 220 คน

ตอนที่ 4.2 การเก็บรวบรวมข้อมูลเพื่อนำมาศึกษาคุณภาพด้านความเที่ยงเมื่อแทนค่าด้วยสัมประสิทธิ์การสรุปอ้างอิง คุณภาพด้านค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ และความตรง เชิงโครงสร้าง

1. นำหนังสือขอความร่วมมือจากบัณฑิตวิทยาลัย ขออนุญาตเก็บข้อมูลในโรงเรียนที่ผู้วิจัยคัดเลือกเป็นกลุ่มตัวอย่าง

2. ดำเนินการวิจัยโดยผู้วิจัยขอความร่วมมือในการประเมินความคิดรวบยอดโดยใช้แผนผัง มโนทัศน์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 จำนวน 220 คน

3. ติดต่อประสานงานและขออนุญาตผู้บริหารโรงเรียนและครูผู้ประเมินที่ใช้เป็นกลุ่มตัวอย่างเพื่อกำหนดวันและเวลา ในการประเมิน

4. ผู้วิจัยประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ ในชั้นมัธยมศึกษาปีที่ 1 ซึ่งทำการประเมินในเรื่อง สมบัติของจำนวนนับ และ ระบบจำนวนเต็ม ตามลำดับ โดยทำการประเมินนักเรียนเมื่อเรียนสิ้นสุดการเรียนในเรื่องดังกล่าว เรื่องละ 1 แผนผัง รวมทั้งสิ้น 2 แผนผัง

5. ผู้วิจัยนำแผนผังมโนทัศน์แบบสร้างผังจากคำที่นักเรียนสร้างขึ้นจำนวน 220 คน ซึ่งนักเรียนคนหนึ่งมีแผนผังมโนทัศน์จำนวน 2 แผนผัง รวมทั้งหมด 440 ผลงาน นำผลงานของนักเรียนมาใส่เลขที่ใหม่ แล้วถ่ายเอกสารเป็น 4 ชุด เพื่อจัดให้แก่ผู้ประเมินดำเนินการให้คะแนน โดยที่ผู้ประเมินกลุ่มที่ 1 จำนวน 2 คน ใช้วิธีการตรวจให้คะแนนของโนแวคและโกวิน และผู้ประเมินกลุ่มที่ 2 จำนวน 2 คน ใช้วิธีการตรวจให้คะแนนของบอลท์

6. ชี้แจงรายละเอียดการใช้วิธีการให้คะแนนทั้ง 2 แบบให้ผู้ประเมินอีกครั้งเข้าใจก่อนเริ่มทำการประเมิน

7. ผู้ประเมินทำการตรวจให้คะแนนตามวิธีให้คะแนนแต่ละแผนผังทีละแผนผังจนครบทุกคน

8. บันทึกคะแนนลงในแบบบันทึกคะแนน

ตอนที่ 4.3 การเก็บรวบรวมข้อมูลเพื่อนำมาศึกษาคุณภาพด้านความตรงตามสภาพ

1. นำหนังสือขอความร่วมมือจากบัณฑิตวิทยาลัย ขออนุญาตเก็บข้อมูลในโรงเรียนที่ผู้วิจัยคัดเลือกเป็นกลุ่มตัวอย่าง

2. ดำเนินการวิจัยโดยผู้วิจัยขอความร่วมมือในการประเมินความคิดรวบยอดวิชาคณิตศาสตร์ โดยใช้แบบทดสอบประเภทเลือกตอบ ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 จำนวน 220 คน

3. ติดต่อประสานงานและขออนุญาตผู้บริหารโรงเรียนและครูผู้ประเมินที่ใช้เป็นกลุ่มตัวอย่างเพื่อกำหนดวันและเวลา ในการประเมิน

4. ผู้วิจัยประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แบบทดสอบประเภทเลือกตอบในชั้นมัธยมศึกษาปีที่ 1 ซึ่งทำการทดสอบใน เรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็มตามลำดับ โดยทำการทดสอบนักเรียนเมื่อเรียนสิ้นสุดการเรียนในเรื่องดังกล่าว เรื่องละ 1 แบบสอบ รวมทั้งสิ้น 2 แบบสอบ

5. ผู้วิจัยนำกระดาษคำตอบของเรียนจำนวน 220 คน มาทำการตรวจให้คะแนนแบบทวิภาค แล้วบันทึกคะแนนลงในแบบบันทึกคะแนน

ตอนที่ 5 การวิเคราะห์ข้อมูล

ผู้วิจัยนำข้อมูลที่ได้ มาวิเคราะห์หาค่าต่างๆโดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป SPSS for window version 10.0 โปรแกรม GENOVA 3.1 และ MULTILOG ในการวิเคราะห์ข้อมูล ผู้วิจัยนำเสนอการวิเคราะห์ข้อมูลเป็น 6 ตอน ดังนี้

ตอนที่ 1 การวิเคราะห์ข้อมูลเบื้องต้น โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป SPSS for window version 10.0

ตอนที่ 2 การวิเคราะห์ความตรงตามสภาพ จากค่าสัมประสิทธิ์สหสัมพันธ์(r)และแปลงค่าความตรงตามสภาพจากค่าสัมประสิทธิ์สหสัมพันธ์ เป็นคะแนนฟิชเชอร์ซี (Z_r)โดยใช้สูตร (Glass and Stanley , 1970)

$$Z_r = \log_e \sqrt{(1+r)/(1-r)}$$

เปรียบเทียบความเป็นอิสระของความตรงตามสภาพในภาพรวม โดยใช้สถิติทดสอบไคว์สแควร์ (χ^2) (Glass and Stanley , 1970) มีสูตร คือ

$$\chi^2_{j-1;\alpha} = \sum (Z_{rj} - Z_r)^2 (n_j - 3)$$

$$Z_r = \frac{\sum (n_j - 3) Z_{rj}}{\sum (n_j - 3)}$$

เมื่อ Z_{rj} แทน คะแนนฟิชเชอร์ซี ของกลุ่มผู้สอบj

n_j แทน จำนวนผู้สอบกลุ่มที่ j

ตอนที่ 3 วิเคราะห์ความตรงเชิงโครงสร้าง (construct validity) ด้วยเมทริกซ์พหุลักษณะพหุวิธี(multitrait-multimethod : MTMM) โดยอาศัยหลักการหาค่าสัมประสิทธิ์สหสัมพันธ์เชิงเส้นตรง จากผลการวัดคุณลักษณะหลายอย่างด้วยวิธีการวัดหลายวิธี โดยวัดความคิดรวบยอดโดยใช้แผนผังมโนทัศน์ สมบัติของจำนวนนับและระบบจำนวนเต็ม และแสดงค่าสัมประสิทธิ์สหสัมพันธ์เหล่านี้ในรูปของเมทริกซ์สหสัมพันธ์

ตอนที่ 4 วิเคราะห์ค่าฟังก์ชันสารสนเทศของแบบสอบ (item information function) และค่าฟังก์ชันสารสนเทศของแบบสอบ (test information function) ตามทฤษฎีการตอบสนองข้อสอบ (item response theory) แบบพหุวิภาคที่ใช้โมเดล GRM (Graded Response Model) โดยใช้โปรแกรม MULTILOG

ตอนที่ 5 การวิเคราะห์คุณภาพด้านความเที่ยงโดยการวิเคราะห์ค่าสัมประสิทธิ์สรุปอ้างอิง (Generalizability Coefficient) โดยใช้โปรแกรม GENOVA 2.2 (GENeralized analysis Of VAriance System) ออกแบบการวัดเป็น two-facet design โดยมีรูปแบบเป็น $P \times I \times R$ crossed design กล่าวคือ (P) แทนจำนวนนักเรียน (I) แทนจำนวนแผนผังมโนทัศน์ และ (R) แทนจำนวนผู้ประเมิน ซึ่งการออกแบบการวัดครั้งนี้ มีผู้ประเมิน 2 คน ตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำของนักเรียนทุกคน โดยใช้วิธีการตรวจให้คะแนนแบบของโนแวคและโกวิน (Novak & Gowin, 1984) และผู้ประเมิน 2 คน ตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำของนักเรียนทุกคน โดยใช้วิธีการตรวจให้คะแนนแบบของบอลท์ (Bolte, 1999)

ตอนที่ 6 วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงเพื่อการตัดสินใจเชิงสรุปอ้างอิง (D-STUDY) เลือกจำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสม โดยแสดงการเปรียบเทียบค่าค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้วิธีการตรวจให้คะแนนของโนแวคและโกวิน และวิธีการตรวจให้คะแนนของบอลท์ กรณีมีผู้ประเมิน 1-6 คน จำนวนแผนผัง 4-28 แผนผัง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาและพัฒนาเกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ด้วยวิธีของโนแวคและโกวิน และวิธีของบอลท์ และเพื่อเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ เมื่อตรวจโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และใช้วิธีการให้คะแนนของบอลท์ โดยพิจารณาจากความตรงตามสภาพ ความตรงเชิงโครงสร้าง ความเที่ยงซึ่งแทนด้วยสัมประสิทธิ์สรุปอ้างอิง ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ และวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสม ซึ่งการวิจัยครั้งในแต่ละวิธีจะมีผู้ประเมิน 2 คน ตรวจให้คะแนนนักเรียน 220 คน คนละ 2 แผนผัง ซึ่งผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลเป็น 3 ตอน ตามลำดับดังนี้

ตอนที่ 1 ผลการศึกษาและพัฒนาเกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์

ตอนที่ 2 ผลการเปรียบเทียบคุณภาพด้านความตรงและความเที่ยง

2.1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

2.2 ผลการเปรียบเทียบความตรงตามสภาพระหว่างวิธีการตรวจให้คะแนนที่แตกต่างกัน

2.3 ผลการตรวจสอบความตรงเชิงโครงสร้างโดยใช้เมทริกซ์พหุลักษณะพหุวิธี

2.4 ผลการเปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบ ฟังก์ชันสารสนเทศของแบบสอบและประสิทธิภาพสัมพัทธ์เฉลี่ยระหว่างวิธีการตรวจให้คะแนนที่แตกต่างกัน

2.5 การวิเคราะห์คุณภาพด้านความเที่ยงโดยการวิเคราะห์ค่าสัมประสิทธิ์สรุปอ้างอิง

ตอนที่ 3 วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงเพื่อการตัดสินใจเชิงสรุปอ้างอิง

ตอนที่ 1 ผลการศึกษาและพัฒนาเกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์

เกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ได้พัฒนาแนวคิดมาจากงานวิจัยของโนแวคและโกวิน (Novak & Gowin, 1984) และงานวิจัยของบอลท์ (Bolte, 1999) ซึ่งมีรายละเอียดดังนี้

เกณฑ์การให้คะแนนของโนแวคและโกวิน จะพิจารณาแผนผังมโนทัศน์ในด้านคุณภาพ 4 ด้าน คือ พิจารณาจากประพจน์ (propositions) การจัดลำดับขั้นตอน (hierarchy) การเชื่อมระหว่างมโนทัศน์ (cross link) และตัวอย่าง (examples) ซึ่งแต่ละด้านจะมีคะแนนเต็ม 4 คะแนน

เกณฑ์การให้คะแนนของบอลท์ จะพิจารณาให้คะแนนในด้านการจัดแผนผังมโนทัศน์ (organization) 6 คะแนน ความถูกต้อง (accuracy) 4 คะแนน

ผู้วิจัยสร้างคำชี้แจงในการใช้เกณฑ์การให้คะแนนและคำอธิบายคุณภาพ และนำเกณฑ์การให้คะแนนไปให้ผู้ทรงคุณวุฒิทางการวัดและประเมินผลการศึกษาจำนวน 3 ท่าน ตรวจสอบความเหมาะสมของวิธีการตรวจให้คะแนน โดยหาค่าดัชนีความสอดคล้องของตัวบ่งชี้ (Item Objective Congruence : IOC) กับคำอธิบายคุณภาพ ดังรายละเอียดในตารางที่ 4.1

ตารางที่ 4.1 ผลการตรวจสอบความตรงตามเนื้อหาโดยวิเคราะห์ค่าดัชนีความสอดคล้อง (IOC) ระหว่างตัวบ่งชี้กับคำคำอธิบายคุณภาพในเกณฑ์การตรวจให้คะแนน ของผู้ทรงคุณวุฒิ 3 ท่าน

ตัวบ่งชี้	ดัชนีความสอดคล้องของความคิดเห็นของผู้ทรงคุณวุฒิ
วิธีการตรวจให้คะแนนของโนแวคและโกวิน	
1. ประพจน์	1.00
2. การจัดลำดับขั้นตอน	0.66
3. การเชื่อมโยงข้ามสาย	0.66
4. ตัวอย่าง	1.00
วิธีการตรวจให้คะแนนของบอลท์	
1. การจัดแผนผังมโนทัศน์	0.66
2. ความถูกต้อง	0.66

จากตารางที่ 4.1 พบว่าผู้ทรงคุณวุฒิมีความเห็นสอดคล้องกัน โดยตัวบ่งชี้กับคำคำอธิบายคุณภาพในเกณฑ์การตรวจให้คะแนน สามารถวัดได้ตรงตามเนื้อหาและมีค่าดัชนีความสอดคล้อง (IOC) ตั้งแต่ 0.66-1.00 ซึ่งถือว่าผู้ทรงคุณวุฒิส่วนใหญ่มีความเห็นสอดคล้องกันมากกว่าร้อยละ 50

ตารางที่ 4.2 เกณฑ์การตรวจให้คะแนนแผนผังมโนทัศน์ของโนแวคและโกวิน

ระดับคุณภาพ	ตัวบ่งชี้
	ประพจน์ (propositions)
4	มโนทัศน์ที่เชื่อมกันตั้งแต่ 75% มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
3	มโนทัศน์ที่เชื่อมกันตั้งแต่ 50% แต่ไม่ถึง 75 % มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
2	มโนทัศน์ที่เชื่อมกันตั้งแต่ 25% แต่ไม่ถึง 50 % มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
1	มโนทัศน์ที่เชื่อมกันไม่ถึง 25 % มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
0	ไม่พบมโนทัศน์ที่เชื่อมกันและคำเชื่อมมีความหมายและตรงกับความสัมพันธ์เลย
ระดับคุณภาพ	ตัวบ่งชี้
	การจัดลำดับขั้นตอน (hierarchy)
4	แผนผังมโนทัศน์มีการจัดลำดับขั้นตอนตั้งแต่ 75% โดยแต่ละมโนทัศน์ที่อยู่รองลงมาเป็นมโนทัศน์ที่เฉพาะเจาะจง และมีความกว้างขวางน้อยกว่ามโนทัศน์ที่อยู่ลำดับแรก ๆ
3	แผนผังมโนทัศน์มีการจัด ลำดับขั้นตอนตั้งแต่ 50% แต่ไม่ถึง 75% โดยมีการจัดลำดับขั้นตอนเป็นส่วนมาก
2	แผนผังมโนทัศน์มีการจัด ลำดับขั้นตอนตั้งแต่ 25% แต่ไม่ถึง 50% โดยมีการจัดลำดับขั้นตอนเป็นบางส่วน
1	แผนผังมโนทัศน์มีการจัดลำดับขั้นตอนไม่ถึง 25% โดยมีการจัดลำดับขั้นตอนน้อยมาก
0	แผนผังมโนทัศน์ไม่มีการมีการจัดลำดับขั้นตอนเลย หรือการจัดลำดับขั้นตอนไม่ถูกต้องเลย
ระดับคุณภาพ	ตัวบ่งชี้
	การเชื่อมระหว่างสายของมโนทัศน์ (cross link)
4	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์ตั้งแต่ 75% อย่างสมเหตุสมผลและถูกต้อง
3	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้องตั้งแต่ 50% แต่ไม่ถึง 75%
2	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้องตั้งแต่ 25% แต่ไม่ถึง 50%
1	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้องไม่ถึง 25%
0	ไม่มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้อง

ตารางที่ 4.2 เกณฑ์การตรวจให้คะแนนแผนผังมโนทัศน์ของโนแวคและโกวิน (ต่อ)

ระดับคุณภาพ	ตัวบ่งชี้
	ตัวอย่าง (example)
4	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องตั้งแต่ 75%
3	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องตั้งแต่ 50% แต่ไม่ถึง 75%
2	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องตั้งแต่ 25% แต่ไม่ถึง 50%
1	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องไม่ถึง 25%
0	ไม่พบการยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้น

ตารางที่ 4.3 เกณฑ์การตรวจให้คะแนนแผนผังมโนทัศน์ของบอลท์

ระดับคุณภาพ	ตัวบ่งชี้
	การจัดแผนผังมโนทัศน์
6	แผนผังมีความสมบูรณ์ โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กันทั้งหมด มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ 100%
5	แผนผังมีความสมบูรณ์เป็นส่วนใหญ่ โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ ได้ตั้งแต่ 80% แต่ไม่ถึง 100%
4	แผนผังมีความถูกต้องสมบูรณ์ปานกลาง มีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ได้ตั้งแต่ 60% แต่ไม่ถึง 80%
3	แผนผังมีความสมบูรณ์เป็นบางส่วน โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ได้ตั้งแต่ 40% แต่ไม่ถึง 80%
2	แผนผังมีความถูกต้องสมบูรณ์เล็กน้อย โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ ได้ตั้งแต่ 20% แต่ไม่ถึง 40%
1	แผนผังมีความถูกต้องสมบูรณ์น้อยมาก โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ได้น้อยกว่า 20%

ตารางที่ 4.3 วิธีการตรวจให้คะแนนแผนผังมโนทัศน์ของบอลท์ (ต่อ)

ระดับคุณภาพ	ตัวบ่งชี้ การจัดแผนผังมโนทัศน์
0	แผนผังไม่มีความถูกต้องสมบูรณ์เลย การเชื่อมโยงมโนทัศน์ที่สัมพันธ์กันไม่ถูกต้อง ไม่ใช้ คำเชื่อมความสัมพันธ์ ไม่ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ ไม่ได้ใช้คำมโนทัศน์ที่กำหนดให้
ระดับคุณภาพ	ตัวบ่งชี้ ความถูกต้อง
4	ไม่ผิดพลาดเลย
3	ผิดในรายละเอียดเล็กน้อย
2	คลาดเคลื่อนเล็กน้อยและเป็นการคลาดเคลื่อนในมโนทัศน์รอง
1	คลาดเคลื่อนในมโนทัศน์หลักเป็นส่วนใหญ่
0	ไม่ถูกต้องเลย หรือคลาดเคลื่อนในมโนทัศน์หลักหลายมโนทัศน์

ตอนที่ 2 ผลการเปรียบเทียบคุณภาพด้านความตรงและความเที่ยง

2.1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

ผลการวิเคราะห์ข้อมูลเบื้องต้น ได้นำเสนอข้อมูลเป็น 2 ส่วน ส่วนแรก คือ ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าคะแนนดิบ (ตารางที่ 4.2) การเปรียบเทียบความแตกต่างของคะแนนเฉลี่ยรายคู่ในแต่ละวิธีการตรวจให้คะแนน (ตารางที่ 4.3) จากการตรวจให้คะแนนแผนผังมโนทัศน์ด้วยวิธีการให้คะแนนของโนแวคและโกวิน และการตรวจให้คะแนนแผนผังมโนทัศน์ด้วยวิธีการให้คะแนนของบอลท์

ตารางที่ 4.4 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจำแนกตามวิธีการตรวจให้คะแนน

วิธีการตรวจให้คะแนน	ผู้ประเมิน	ข้อ	คะแนน					
			Min	Max	Mean (คะแนนเต็ม)	Mean%	S.D.	C.V.
โนแวคและโกวิน	คนที่ 1	1	0	16	9.63 (16)	60.19	2.75	4.57
		2	0	16	9.06 (16)	56.63	2.99	5.28
		เฉลี่ย			9.35 (16)	58.44	2.52	
	คนที่ 2	1	2	16	9.30 (16)	58.13	1.40	2.41
		2	1	16	9.03 (16)	56.44	1.33	2.36
		เฉลี่ย			9.17 (16)	57.31	1.13	
	เฉลี่ย				9.26 (16)	57.88	1.69	2.92
บอลท์	คนที่ 3	1	0	8	3.05 (10)	30.50	1.55	5.08
		2	0	10	3.13 (10)	31.30	1.95	6.23
		เฉลี่ย			3.09 (10)	30.90	1.31	
	คนที่ 4	1	1	8	3.51 (10)	35.10	1.83	5.21
		2	0	10	3.60 (10)	36.00	1.86	5.17
		เฉลี่ย			3.56 (10)	35.60	1.51	
	เฉลี่ย				3.32 (10)	33.20	1.34	4.04

จากตารางที่ 4.4 เมื่อพิจารณาค่าร้อยละของค่าเฉลี่ย (Mean %) พบว่า การให้คะแนนวิธีของโนแวคและโกวินมีค่าเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์ ในขณะที่สัมประสิทธิ์ของการกระจาย (C.V.) การให้คะแนนวิธีของบอลท์ มีค่าสูงกว่าวิธีการให้คะแนนของโนแวคและโกวิน แสดงว่าวิธีการให้คะแนนของบอลท์มีการกระจายของคะแนนมากกว่าวิธีของโนแวคและโกวิน

ตารางที่ 4.5 ค่า t ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนระหว่างวิธีการให้คะแนนที่ต่างกัน

การเปรียบเทียบระหว่าง		Paired Differences		t	p-value
		Mean	SD		
โนแวคและโกวิน	บอลท์	24.617	14.76	24.737**	0.000

** p-value < 0.05

จากตารางที่ 4.5 พบว่าวิธีการให้คะแนนของโนแวคและโกวินมีคะแนนเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์ และเมื่อเปรียบเทียบความแตกต่างของค่าเฉลี่ยเป็นรายคู่พบว่าวิธีการให้คะแนนทั้ง 2 วิธี มีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

2.2 ผลการเปรียบเทียบความตรงตามสภาพ ระหว่างวิธีการตรวจให้คะแนนที่แตกต่างกัน

การคำนวณความตรงตามสภาพ ของแบบสอบที่มีวิธีการตรวจให้คะแนนที่แตกต่างกัน พิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ (r) ระหว่างคะแนนของผู้สอบที่ได้จากการสอบด้วยแบบสอบประเภทเลือกตอบประกอบด้วย 2 เรื่อง คือ สมบัติของจำนวนนับ และระบบจำนวนเต็ม ที่ผู้วิจัยสร้างขึ้น กับคะแนนที่ได้จากการประเมินแผนผังมโนทัศน์ โดยที่การตรวจให้คะแนนจากการประเมินแผนผังมโนทัศน์มีการตรวจให้คะแนน 2 วิธีที่แตกต่าง และมีการแปลงค่าสัมประสิทธิ์สหสัมพันธ์ให้เป็นคะแนนฟิชเชอร์ซี (Fisher's Z) จากสูตรของ Glass and Stanley (1970) เพื่อนำมาเปรียบเทียบความแตกต่างของค่าความตรงตามสภาพในภาพรวมโดยใช้สถิติทดสอบไคว์สแควร์ (χ^2) ได้ผลการวิเคราะห์แสดงในตารางที่ 4.6 และ 4.7

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.6 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากแบบทดสอบประเภทเลือกตอบ

เนื้อหา	n	Min	Max	Mean	S.D.
แบบทดสอบประเภทเลือกตอบ					
สมบัติของจำนวนนับ	220	1	13	6.635	2.579
ระบบจำนวนเต็ม	220	1	13	5.977	1.996
เฉลี่ย	220	1	13	6.307	1.949

จากตารางที่ 4.6 พบว่าค่าเฉลี่ยที่ได้จากแบบทดสอบประเภทเลือกตอบในเรื่องสมบัติของจำนวนนับ และ ระบบจำนวนเต็ม มีค่าเท่ากับ 6.635 และ 5.977 ตามลำดับ และส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 2.579 และ 1.996 ตามลำดับ

ตารางที่ 4.7 ความตรงตามสภาพจากการสอบจำแนกตามวิธีการตรวจให้คะแนน

วิธีการตรวจให้คะแนน	จำนวนผู้สอบ	ความตรงตามสภาพ		χ^2
		r	Z	
โนแวคและโกวิน (Novak&Gowin)	220	.829**	1.185	8.99**
บอลท์ (Bolte)	220	.715**	0.897	

** p-value < 0.05

จากตารางที่ 4.7 พบว่าความตรงตามสภาพที่พิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ (r) ระหว่างการตรวจให้คะแนน 2 วิธี กับคะแนนที่ได้จากแบบสอบประเภทเลือกตอบมีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.829และ 0.715 และเมื่อนำมาเปรียบเทียบความแตกต่างโดยใช้คะแนนฟิชเชอร์ซี เปรียบเทียบในภาพรวมโดยใช้สถิติทดสอบไคว์สแควร์ (χ^2) พบว่าวิธีการตรวจให้คะแนนทั้ง 2 วิธี มีความตรงตามสภาพแตกต่างกัน แสดงว่าการตรวจด้วยวิธีการให้คะแนนของโนแวคและโกวินให้ความตรงตามสภาพแตกต่างกับการตรวจด้วยวิธีการให้คะแนนของบอลท์

2.3 ผลการตรวจสอบความตรงเชิงโครงสร้างโดยใช้เมทริกซ์พหุลักษณะพหุวิธี (Multitrait-Multimethod Matrix)

ผลการตรวจสอบความตรงของการตรวจให้คะแนนแผนผังมโนทัศน์ด้วยวิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ ในส่วนนี้จะเป็นการตรวจสอบความตรงเชิงโครงสร้างโดยใช้เมทริกซ์พหุลักษณะพหุวิธี (Multitrait-Multimethod Matrix) ประกอบด้วยการวัดคุณลักษณะ 2 อย่าง (สมบัติของจำนวนนับ และ ระบบจำนวนเต็ม) ด้วยวิธีการวัด 3 วิธี (แบบทดสอบประเภทเลือกตอบ, แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และแผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของบอลท์) ดังนั้น เมทริกซ์พหุลักษณะพหุวิธีในการศึกษานี้จะมีขนาด 3x2 โดยมีรายละเอียดของการวิเคราะห์ดังตารางที่ 4.8 และ 4.9

ตารางที่ 4.8 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบจากแบบทดสอบประเภทเลือกตอบ แผนผังมโนทัศน์ที่ตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน และตรวจให้คะแนนด้วยวิธีของบอลท์

เรื่อง	n	Min	Max	Mean	S.D.
แบบทดสอบประเภทเลือกตอบ					
สมบัติของจำนวนนับ	220	1	13	6.635	2.579
ระบบจำนวนเต็ม	220	1	13	5.977	1.996
วิธีตรวจให้คะแนนของโนแวคและโกวิน					
สมบัติของจำนวนนับ	220	1	15.50	9.466	1.883
ระบบจำนวนเต็ม	220	0.05	15.50	9.048	1.962
วิธีตรวจให้คะแนนของบอลท์					
สมบัติของจำนวนนับ	220	1	8	3.282	1.621
ระบบจำนวนเต็ม	220	0	10	3.366	1.825

จากตารางที่ 4.8 พบว่า โดยภาพรวมแล้ว เรื่องสมบัติของจำนวนนับให้คะแนนสูงกว่าระบบจำนวนเต็ม โดยที่นักเรียนทำแบบทดสอบประเภทเลือกตอบและการตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำด้วยวิธีของโนแวคและโกวินในเรื่องสมบัติของจำนวนนับได้ดีกว่า

เรื่องระบบจำนวนเต็ม (Mean = 6.635 vs. Mean = 5.977) , (Mean = 9.466 vs. Mean = 9.048) สำหรับคะแนนจากการตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำด้วยวิธีของบอลท์ในเรื่องระบบจำนวนเต็มมีค่าสูงกว่าเรื่องสมบัติของจำนวนนับ (Mean = 3.366 vs. Mean = 3.282)

ตารางที่ 4.9 เมทริกซ์พหุลักษณะพหุวิธี

		MC		**CmNG**	
		A1	**B1**	**A2**	**B2**
MC	**A1**	(0.467)			
	B1	0.444	(0.498)		
CmNG	**A2**	0.768	0.312	(0.412)	
	B2	0.289	0.817	0.540	(0.453)
		MC		**CmB**	
		A1	**B1**	**A3**	**B3**
MC	**A1**	(0.467)			
	B1	0.444	(0.498)		
CmB	**A3**	0.704	0.148	(0.266)	
	B3	0.257	0.781	0.208	(0.360)

MC = แบบทดสอบประเภทเลือกตอบ

CmNG = แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน

CmB = แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของบอลท์

A = สมบัติของจำนวนนับ

B = ระบบจำนวนเต็ม

จากตารางที่ 4.9 สำหรับการวัดด้วยแบบทดสอบประเภทเลือกตอบกับแผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน พบว่าการวัดคุณลักษณะเดียวที่วัดด้วยวิธีต่างกัน (A1-A2 และ B1-B2) จะให้ค่าสหสัมพันธ์เท่ากับ 0.768 และ 0.817 และการวัดคุณลักษณะต่างกันวัดด้วยวิธีเดียวกัน (A1-B1 และ A2-B2) จะมีค่าสหสัมพันธ์เท่ากับ 0.444 และ 0.540 และการวัดคุณลักษณะต่างกันด้วยวิธีต่างกัน (A1-B2 และ B1-A2) จะให้ค่าสหสัมพันธ์เท่ากับ 0.289 และ 0.312

และ การวัดด้วยแบบทดสอบประเภทเลือกตอบกับแผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของบอลท์ พบว่าการวัดคุณลักษณะเดียวที่วัดด้วยวิธีต่างกัน (A1-A3 และ B1B3) จะให้ค่าสหสัมพันธ์เท่ากับ 0.704 และ 0.781 และการวัดคุณลักษณะต่างกันวัดด้วยวิธีเดียวกัน (A1-B1 และ A3-B3) จะมีค่าสหสัมพันธ์เท่ากับ 0.444 และ 0.208 และการวัดคุณลักษณะต่างกันด้วยวิธีต่างกัน (A1-B3 และ B1-A3) จะให้ค่าสหสัมพันธ์เท่ากับ 0.257 และ 0.148

ดังนั้นสรุปได้ว่า แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และแผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของบอลท์มีความตรงเชิงโครงสร้าง

2.4 ผลการเปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบ ฟังก์ชันสารสนเทศของแบบสอบ และอัตราส่วนสารสนเทศระหว่างวิธีการตรวจให้คะแนนที่แตกต่างกัน

ผลการวิเคราะห์เปรียบเทียบในตอนนี้ ได้นำเสนอ เป็น 2 ส่วน คือ

2.4.1 ผลการเปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบตามวิธีการตรวจให้คะแนนที่แตกต่างกัน

2.4.2 ผลการเปรียบเทียบฟังก์ชันสารสนเทศของแบบสอบ และอัตราส่วนสารสนเทศตามวิธีการตรวจให้คะแนนที่แตกต่างกัน

โดยมีรายละเอียดของผลการวิเคราะห์ ดังนี้

2.4.1 ผลการเปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบ ตามวิธีการตรวจให้คะแนนที่แตกต่างกัน

ผลการวิเคราะห์ค่าฟังก์ชันสารสนเทศของข้อสอบตามวิธีการตรวจให้คะแนนที่แตกต่างกัน โดยวิเคราะห์จากคะแนนที่ได้จากการตอบและตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำ โดยการตรวจให้คะแนนของโนแวคและโกวิน ให้คะแนนเป็นลำดับขั้น ตั้งแต่ขั้นต่ำสุดถึงสูงสุด เป็น 0, 1, 2, 3, 4 ซึ่งมี 4 ลำดับขั้น ส่วนวิธีการให้คะแนนของบอลท์ ให้คะแนนเป็นลำดับขั้นตั้งแต่ขั้นต่ำสุดถึงสูงสุด เป็น 0, 1, 2, 3, 4 , 5, 6 ซึ่งมี 6 ลำดับขั้น

การเปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบจำแนกตามคะแนนความสามารถของผู้สอบ (7 ระดับความสามารถ) แสดงในตารางที่ 4.10- 4.14

ตารางที่ 4.10 ฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยที่ได้จากการตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน จำแนกตามผู้ประเมินและความสามารถของผู้สอบ

ความสามารถ (θ)	ผู้ประเมิน	ฟังก์ชันสารสนเทศของข้อสอบ			
		ข้อ 1	ข้อ 2	Mean	S.D.
-3.00	ผู้ประเมินคนที่ 1	1.720	0.833	1.277	0.627
	ผู้ประเมินคนที่ 2	0.382	0.096	0.239	0.202
-2.00	ผู้ประเมินคนที่ 1	1.794	0.959	1.377	0.590
	ผู้ประเมินคนที่ 2	0.384	0.507	0.446	0.087
-1.00	ผู้ประเมินคนที่ 1	1.459	0.910	1.185	0.388
	ผู้ประเมินคนที่ 2	0.378	1.535	0.957	0.818
0.00	ผู้ประเมินคนที่ 1	1.817	0.940	1.379	0.620
	ผู้ประเมินคนที่ 2	0.301	1.258	0.780	0.677
1.00	ผู้ประเมินคนที่ 1	1.540	0.975	1.258	0.400
	ผู้ประเมินคนที่ 2	0.351	0.920	0.636	0.402
2.00	ผู้ประเมินคนที่ 1	1.733	0.939	1.336	0.561
	ผู้ประเมินคนที่ 2	0.386	1.386	0.886	0.707
3.00	ผู้ประเมินคนที่ 1	0.796	0.454	0.625	0.242
	ผู้ประเมินคนที่ 2	0.331	1.187	0.759	0.605

จากตารางที่ 4.10 พบว่า เมื่อเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบจำแนกตามค่าความสามารถของผู้สอบและผู้ประเมิน เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน พบว่า ผู้ประเมินคนที่ 1 ตรวจแผนผังมโนทัศน์แบบสร้างผังจากคำให้ค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยสูงที่สุดในระดับความสามารถ 0 และมีค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเท่ากับ 1.379 ในขณะที่ผู้ประเมินคนที่ 2 ตรวจแผนผังมโนทัศน์แบบสร้างผังจากคำให้ค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยสูงที่สุดในระดับความสามารถ -1 มีค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเท่ากับ 0.957

ตารางที่ 4.11 เปรียบเทียบความแตกต่างของค่าเฉลี่ยของฟังก์ชันสารสนเทศของข้อสอบ เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน โดยจำแนกตามผู้ประเมิน

ความสามารถ (θ)	ผู้ประเมิน	Paired Differences		t-test	p-value
		Mean	S.D.		
-3	ผู้ประเมินคนที่ 1 ผู้ประเมินคนที่ 2	1.038	0.425	3.453	0.179
-2	ผู้ประเมินคนที่ 1 ผู้ประเมินคนที่ 2	0.931	0.677	1.944	0.303
-1	ผู้ประเมินคนที่ 1 ผู้ประเมินคนที่ 2	0.228	1.206	0.267	0.834
0	ผู้ประเมินคนที่ 1 ผู้ประเมินคนที่ 2	0.599	1.297	0.653	0.632
1	ผู้ประเมินคนที่ 1 ผู้ประเมินคนที่ 2	0.622	0.802	1.097	0.471
2	ผู้ประเมินคนที่ 1 ผู้ประเมินคนที่ 2	0.450	1.269	0.502	0.704
3	ผู้ประเมินคนที่ 1 ผู้ประเมินคนที่ 2	0.134	0.847	0.224	0.860

** p-value < 0.05

จากตารางที่ 4.11 พบว่า เมื่อเปรียบเทียบความแตกต่างของค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเป็นรายคู่พบว่าผู้ประเมินทั้ง 2 คน ให้ค่าฟังชันสารสนเทศของข้อสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ0.05 ในทุกระดับความสามารถ

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.12 ฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยที่ได้จากการตรวจให้คะแนนด้วยวิธีของบอลท์ จำแนกตามผู้ประเมินและความสามารถของผู้สอบ

ความสามารถ (θ)	ผู้ประเมิน	ฟังก์ชันสารสนเทศของข้อสอบ			
		ข้อ 1	ข้อ 2	Mean	S.D.
-3.00	ผู้ประเมินคนที่ 3	0.093	0.077	0.085	0.011
	ผู้ประเมินคนที่ 4	0.025	0.144	0.085	0.084
-2.00	ผู้ประเมินคนที่ 3	0.196	0.105	0.151	0.064
	ผู้ประเมินคนที่ 4	0.250	0.270	0.260	0.014
-1.00	ผู้ประเมินคนที่ 3	0.288	0.122	0.205	0.117
	ผู้ประเมินคนที่ 4	1.282	0.330	0.806	0.673
0.00	ผู้ประเมินคนที่ 3	0.287	0.127	0.207	0.113
	ผู้ประเมินคนที่ 4	1.030	0.327	0.679	0.497
1.00	ผู้ประเมินคนที่ 3	0.287	0.129	0.208	0.112
	ผู้ประเมินคนที่ 4	1.608	0.345	0.977	0.893
2.00	ผู้ประเมินคนที่ 3	0.309	0.133	0.221	0.124
	ผู้ประเมินคนที่ 4	1.019	0.340	0.680	0.480
3.00	ผู้ประเมินคนที่ 3	0.291	0.136	0.214	0.110
	ผู้ประเมินคนที่ 4	0.149	0.355	0.252	0.146

จากตารางที่ 4.12 พบว่าเมื่อเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบจำแนกตามค่าความสามารถของผู้สอบและผู้ประเมิน เมื่อตรวจให้คะแนนด้วยวิธีของบอลท์ พบว่า ผู้ประเมินคนที่ 3 ตรวจแผนผังมโนทัศน์แบบสร้างผังจากคำให้ค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยสูงที่สุดในระดับความสามารถ 2 และมีค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเท่ากับ 0.221 ในขณะที่ผู้ประเมินคนที่ 4 ตรวจแผนผังมโนทัศน์แบบสร้างผังจากคำให้ค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยสูงที่สุดในระดับความสามารถ 1 มีค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเท่ากับ 0.977

ตารางที่ 4.13 เปรียบเทียบความแตกต่างของค่าเฉลี่ยของฟังก์ชันสารสนเทศของข้อสอบ เมื่อตรวจให้คะแนนด้วยวิธีของบอลท์ โดยจำแนกตามผู้ประเมิน

ความสามารถ (θ)	ผู้ประเมิน	Paired Differences		t-test	p-value
		Mean	S.D.		
-3	ผู้ประเมินคนที่ 3 ผู้ประเมินคนที่ 4	0.001	0.095	0.007	0.995
-2	ผู้ประเมินคนที่ 3 ผู้ประเมินคนที่ 4	0.110	0.078	1.973	0.299
-1	ผู้ประเมินคนที่ 3 ผู้ประเมินคนที่ 4	0.601	0.556	1.529	0.369
0	ผู้ประเมินคนที่ 3 ผู้ประเมินคนที่ 4	0.472	0.384	1.737	0.333
1	ผู้ประเมินคนที่ 3 ผู้ประเมินคนที่ 4	0.769	0.781	1.391	0.397
2	ผู้ประเมินคนที่ 3 ผู้ประเมินคนที่ 4	0.459	0.356	1.823	0.319
3	ผู้ประเมินคนที่ 3 ผู้ประเมินคนที่ 4	0.039	0.255	0.213	0.866

** p-value < 0.05

จากตารางที่ 4.13 พบว่า เมื่อเปรียบเทียบความแตกต่างของค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเป็นรายคู่พบว่าผู้ประเมินทั้ง 2 คน ให้ค่าฟังชันสารสนเทศของข้อสอบไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ในทุกระดับความสามารถ

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.14 ฟังก์ชันสารสนเทศของข้อสอบเฉลี่ย จำแนกตามความสามารถของผู้สอบและวิธีการตรวจให้คะแนน

ความสามารถ (θ)	วิธีการตรวจให้คะแนน	ฟังก์ชันสารสนเทศของข้อสอบเฉลี่ย			
		Min	Max	Mean	S.D.
-3.00	โนแวคและโกวิน	0.096	1.720	0.758	0.710
	บอลท์	0.025	0.144	0.085	0.049
-2.00	โนแวคและโกวิน	0.384	1.794	0.911	0.638
	บอลท์	0.105	0.270	0.205	0.074
-1.00	โนแวคและโกวิน	0.378	1.535	1.071	0.539
	บอลท์	0.122	1.282	0.309	0.030
0.00	โนแวคและโกวิน	0.301	1.817	1.079	0.633
	บอลท์	0.127	1.030	0.307	0.028
1.00	โนแวคและโกวิน	0.351	1.540	0.947	0.486
	บอลท์	0.129	1.608	0.316	0.041
2.00	โนแวคและโกวิน	0.386	1.733	1.111	0.582
	บอลท์	0.133	1.019	0.325	0.022
3.00	โนแวคและโกวิน	0.331	1.187	0.692	0.384
	บอลท์	0.136	0.355	0.233	0.108

จากตารางที่ 4.14 เมื่อเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบจำแนกตามค่าความสามารถของผู้สอบตามวิธีการตรวจให้คะแนน พบว่า วิธีการตรวจให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ ให้ค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยสูงที่สุดเหมือนกันในระดับความสามารถ 2 และมีค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเท่ากับ 1.111 และ 0.325 ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.15 เปรียบเทียบฟังก์ชันสารสนเทศของข้อสอบโดยเฉลี่ยจำแนกตามวิธีการตรวจให้คะแนน

วิธีตรวจให้คะแนน	Paired Differences		t-test	p-value
	Mean	S.D.		
โนแวคและโกวิน บอลท์	0.684	0.114	15.848**	0.000

** p-value < 0.05

จากตารางที่ 4.15 เมื่อเปรียบเทียบความแตกต่างของค่าฟังก์ชันสารสนเทศของข้อสอบเฉลี่ยเป็นรายคู่ พบว่า ทั้งสองวิธีมีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยที่วิธีการให้คะแนนของโนแวคและโกวินมีค่าฟังก์ชันสารสนเทศของข้อสอบโดยเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์

2.4.2 ผลการเปรียบเทียบฟังก์ชันสารสนเทศของแบบสอบตามวิธีการให้คะแนนที่แตกต่างกัน

การเปรียบเทียบฟังก์ชันสารสนเทศของแบบสอบจำแนกตามวิธีการให้คะแนนทั้ง 2 วิธี แสดงในตารางที่ 4.16

ตารางที่ 4.16 ฟังก์ชันสารสนเทศของแบบสอบ จำแนกตามความวิธีการตรวจให้คะแนน

วิธีการตรวจให้คะแนน	ค่าฟังก์ชันสารสนเทศของแบบสอบที่ระดับ (θ) ต่างๆ						
	-3	-2	-1	0	1	2	3
โนแวคและโกวิน	2.516	2.822	3.142	3.159	2.893	1.722	1.384
บอลท์	1.170	1.411	2.011	1.886	2.185	1.900	1.466

จากตารางที่ 4.17 ผลการวิเคราะห์ที่ระดับ θ = -3 ถึง 3 พบว่าวิธีการให้คะแนนของโนแวคและโกวินให้ค่าฟังชันสารสนเทศสูงกว่าวิธีการให้คะแนนนของบอลท์ ในทุกระดับความสามารถ

ตารางที่ 4.17 ค่าสารสนเทศเฉลี่ยของแบบสอบ จำแนกตามวิธีการตรวจให้คะแนน

วิธีการตรวจให้คะแนน	จำนวนผู้สอบ	ค่าสารสนเทศเฉลี่ยของแบบสอบ
โนแวคและโกวิน	220	17.636
บอลท์	220	12.027

จากตารางที่ 4.17 พบว่า วิธีการให้คะแนนของโนแวคและโกวิน ให้ค่าสารสนเทศเฉลี่ยของแบบสอบสูงกว่าวิธีการให้คะแนนนของบอลท์

ตารางที่ 4.18 อัตราส่วนสารสนเทศเฉลี่ยจำแนกตามวิธีการตรวจให้คะแนน

เปรียบเทียบระหว่าง		อัตราส่วนสารสนเทศเฉลี่ย
โนแวคและโกวิน	บอลท์	1.466

จากตารางที่ 4.18 พบว่าอัตราส่วนสารสนเทศเฉลี่ยระหว่างวิธีการให้คะแนนของโนแวคและโกวินและวิธีการให้คะแนนของบอลท์ มีค่าเท่ากับ 1.466 แสดงให้เห็นว่าวิธีการให้คะแนนของโนแวคและโกวินมีประสิทธิภาพสัมพัทธ์สูงกว่าวิธีการให้คะแนนของบอลท์

2.5 การวิเคราะห์คุณภาพด้านความเที่ยงโดยการวิเคราะห์ค่าสัมประสิทธิ์สรุปอ้างอิง (Generalizability Coefficient)

การวิเคราะห์คุณภาพด้านความเที่ยงโดยการวิเคราะห์ค่าสัมประสิทธิ์สรุปอ้างอิง (Generalizability Coefficient) แบ่งการนำเสนอเป็น 2 ตอน ดังนี้

2.5.1 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมิน จำแนกตามผู้ประเมินและตามรูปแบบวิธีการให้คะแนน

2.5.2 การหาความเที่ยงของคะแนนการประเมินแผนผังมโนทัศน์ในแต่ละรูปแบบวิธีการให้คะแนนจากค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient)

2.5.1 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมิน จำแนกตามผู้ประเมินและตามรูปแบบวิธีการให้คะแนน

ผลการวิเคราะห์ข้อมูลตามตารางที่ 4.19 พบว่า ค่าสถิติพื้นฐานของคะแนนที่ได้จากการตรวจให้คะแนนของผู้ประเมินแต่ละคนเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวินมีดังนี้ ค่าเฉลี่ยเท่ากับ 9.35 และ 9.17 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 2.52 และ 1.13 ตามลำดับ และค่าเฉลี่ยรวมมีค่าเท่ากับ 9.26 ส่วนเบี่ยงเบนมาตรฐานรวมมีค่าเท่ากับ 1.83

และเมื่อใช้วิธีการให้คะแนนของบอลท์ มีค่าเฉลี่ยเท่ากับ 3.06 และ 4.01 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 1.27 และ 1.51 ตามลำดับ และค่าเฉลี่ยรวมมีค่าเท่ากับ 3.54 ส่วนเบี่ยงเบนมาตรฐานรวมมีค่าเท่ากับ 1.39

และจากการทดสอบความแตกต่างของค่าเฉลี่ยของการให้คะแนนระหว่างผู้ประเมินพบว่า เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินนั้นให้คะแนนเฉลี่ยของผู้ประเมินทั้งสองคนมีค่าไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 (t=1.33, p=0.185) ในขณะที่การตรวจให้คะแนนด้วยวิธีบอลท์ให้คะแนนเฉลี่ยของผู้ประเมินทั้งสองคนมีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 (t=7.996, p=0.000) โดยที่ผู้ประเมินคนที่ 4 ตรวจให้คะแนนมีค่าเฉลี่ยสูงกว่าผู้ประเมินคนที่ 3

ตารางที่ 4.19 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมิน จำแนกตามผู้ประเมินและตามรูปแบบวิธีการให้คะแนน

รูปแบบการให้คะแนน	ผู้ประเมินคนที่	คะแนนเต็ม	$\bar{X}$	SD	t	p-value
โนแวคและโกวิน	1	16	9.35	2.52	1.330	0.185
	2	16	9.17	1.13		
บอลท์	3	10	3.09	1.27	7.966**	0.000
	4	10	3.56	1.51		

** p-value < 0.05

2.5.2 การหาความเที่ยงของคะแนนการประเมินแผนผังมโนทัศน์ในแต่ละรูปแบบวิธีการให้คะแนนจากค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient)

จุดมุ่งหมายในการวิเคราะห์ข้อมูลตอนนี้ เพื่อคำนวณหาความแปรปรวนตามแหล่งความแปรปรวนต่างๆในเอกภพค่าสังเกตที่เป็นไปได้ทั้งหมดที่แยกเป็นอิสระจากกันขององค์ประกอบ 3 องค์ประกอบ ได้แก่ นักเรียน (P) แผนผังมโนทัศน์ (I) และผู้ประเมิน (R) รวมทั้งผลร่วมระหว่างนักเรียนกับแผนผังมโนทัศน์ (PI) ผลร่วมระหว่างนักเรียนกับผู้ประเมิน (PR) และส่วนที่เหลือ (PIR) ของคะแนนการประเมินแผนผังมโนทัศน์ เมื่อใช้รูปแบบวิธีการให้คะแนนในแต่ละรูปแบบ

ตารางที่ 4.20 ผลการเปรียบเทียบค่าความแปรปรวนและค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์เมื่อใช้รูปแบบวิธีการให้คะแนนของโนแวคและโกวิน (Novak & Gowin) ($n_p = 220,\ n_i = 2,\ n_r = 2$)

แหล่งความแปรปรวน	df	วิธีการตรวจให้คะแนน			
		โนแวคและโกวิน		บอลท์	
		การประมาณค่าความแปรปรวน	ร้อยละของความแปรปรวนทั้งหมด	การประมาณค่าความแปรปรวน	ร้อยละของความแปรปรวนทั้งหมด
P	219	0.13135	25.8859	0.0474	8.12061
I	1	0.00174	0.3429	0.0879	15.0591
R	1	0.00000	0.0000	0.0153	2.62121
PI	219	0.00182	0.3587	0.0057	0.97652
PR	219	0.00000	0.0000	0.0000	0.00000
IR	1	0.01091	2.1500	0.1912	32.7566
PIR	219	0.36160	71.2625	0.2362	40.4660
TOTAL	879	0.50742	100.00	0.5837	100.00
ρ^2_{Abs}		0.81410		0.71484	

เมื่อ P แทน นักเรียน
I แทน แผนผังมโนทัศน์
R แทน ผู้ประเมิน
PIR แทน ส่วนที่เหลือ
PI แทน นักเรียนกับแผนผังมโนทัศน์
PR แทน นักเรียนกับผู้ประเมิน
IR แทน แผนผังมโนทัศน์กับผู้ประเมิน

ผลการวิเคราะห์ข้อมูลตามตารางที่ 4.20 พบว่า แผนผังมโนทัศน์แบบสร้างผังจากคำ เมื่อตรวจให้คะแนนด้วยวิธีที่แตกต่างกัน 2 วิธี ให้ค่าสัมประสิทธิ์การสรุปอ้างอิงแตกต่างกัน โดยที่ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์เมื่อใช้วิธีการให้คะแนนของโนแวคและโกวินมีค่าเท่ากับ 0.81410 และค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์เมื่อใช้วิธีการให้คะแนนของบอลท์มีค่าเท่ากับ 0.71484 ทั้งนี้อาจเป็นเพราะองค์ประกอบของความแปรปรวนของนักเรียนมีค่าสูงในวิธีการให้คะแนนของโนแวคและโกวินมากกว่าวิธีการให้คะแนนของบอลท์

ตอนที่ 3 วิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงเพื่อการตัดสินใจเชิงสรุปอ้างอิง (D-STUDY)

การวิเคราะห์ขั้นตอนนี้เป็นการใช้ข้อมูลจากการศึกษาเพื่อการสรุปอ้างอิงมาศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-STUDY) ซึ่งเป็นการลดความคลาดเคลื่อนและเพิ่มความแม่นยำในการสรุปอ้างอิงโดยมีเงื่อนไขการวัดคือ จำนวนแผนผังมโนทัศน์และผู้ประเมิน ให้จำนวนแผนผังอยู่ระหว่าง 4-28 แผนผัง และจำนวนผู้ประเมิน 1-6 คน ซึ่งเป็นจำนวนที่อยู่ในวิสัยที่สามารถนำไปปฏิบัติได้ โดยการตัดสินใจการวิเคราะห์ว่า จำนวนแผนผังมโนทัศน์กี่แผนผังและจำนวนผู้ประเมินเท่าใดให้ค่าสัมประสิทธิ์การสรุปอ้างอิงสูงสุด เพื่อให้เห็นการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงที่ได้จากการตรวจให้คะแนนทั้ง 2 รูปแบบอย่างชัดเจน ผู้วิจัยได้นำเสนอปรากฏผลดังตาราง

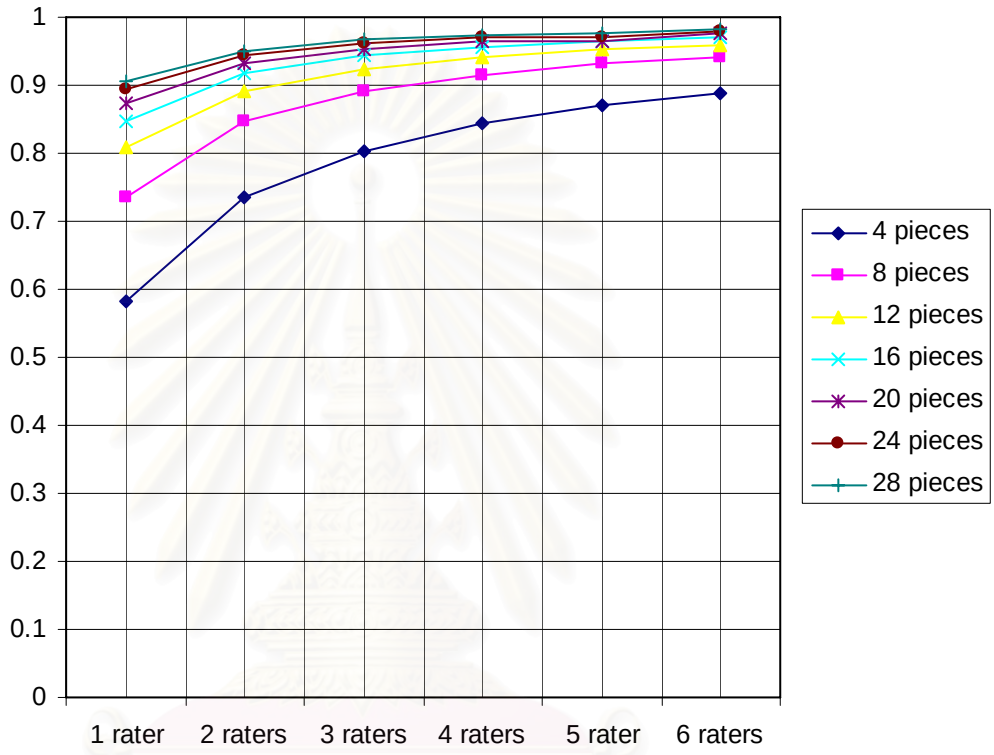
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.21 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้วิธีการตรวจให้คะแนนของโนแวคและโกวิน และวิธีการตรวจให้คะแนนของบอลท์ กรณีมีผู้ประเมิน 1-6 คน (n_p = 220, n_i = 4-28, n_r = 1-6)

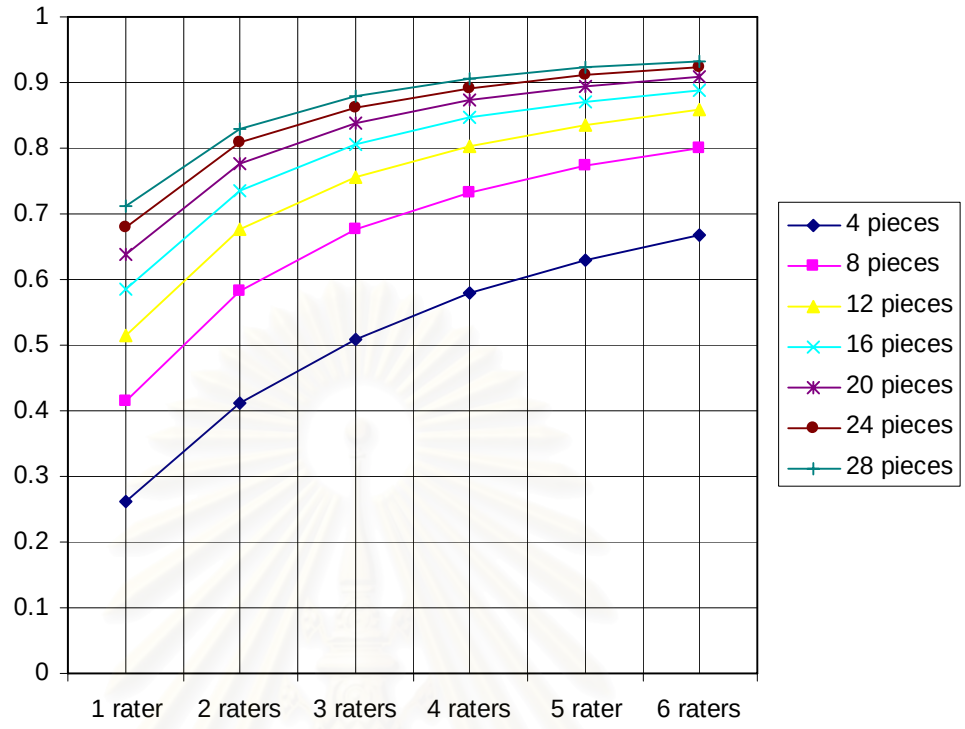
วิธีการให้คะแนน	จำนวนแผนผังมโนทัศน์ (n_i)								
	n_r		4	8	12	16	20	24	28
โนแวคและโกวิน	1	ρ^2_{Abs}	0.5828	0.7364	0.80737	0.8482	0.8748	0.8934	0.9072
	2	ρ^2_{Abs}	0.7346	0.8470	0.8925	0.91716	0.9326	0.9432	0.9509
	3	ρ^2_{Abs}	0.8044	0.8916	0.9250	0.9427	0.9536	0.9611	0.9664
	4	ρ^2_{Abs}	0.8446	0.9157	0.9422	0.95602	0.9641	0.9702	0.9744
	5	ρ^2_{Abs}	0.8707	0.9309	0.9528	0.9642	0.9642	0.9711	0.9758
	6	ρ^2_{Abs}	0.8889	0.9412	0.9600	0.9697	0.9756	0.9796	0.9825
บอลท์	1	ρ^2_{Abs}	0.2612	0.4142	0.5147	0.5858	0.6387	0.6796	0.7122
	2	ρ^2_{Abs}	0.4116	0.5832	0.6773	0.7367	0.7777	0.8076	0.8304
	3	ρ^2_{Abs}	0.5094	0.6750	0.7570	0.8060	0.8385	0.8617	0.8791
	4	ρ^2_{Abs}	0.5781	0.7326	0.8043	0.8457	0.8726	0.8915	0.9056
	5	ρ^2_{Abs}	0.6289	0.7722	0.8356	0.8714	0.8944	0.9105	0.9223
	6	ρ^2_{Abs}	0.6681	0.8010	0.8579	0.8895	0.9096	0.92353	0.9337

จากตารางที่ 4.21 พบว่าเมื่อกำหนดผู้ประเมินจำนวน 1-6 คน จำนวนแผนผังมโนทัศน์ 4-28 แผนผัง การตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน มีค่าสัมประสิทธิ์การสรุปอ้างอิงอยู่ระหว่าง 0.5828 – 0.9825 และเมื่อตรวจให้คะแนนด้วยวิธีของบอลท์ มีค่าสัมประสิทธิ์อ้างอิงอยู่ระหว่าง 0.2612 – 0.9337 และพบว่าจำนวนผู้ประเมินและจำนวนแผนผังที่เหมาะสม ที่ทำให้ค่าสัมประสิทธิ์สรุปอ้างอิงอยู่ในเกณฑ์เหมาะสม (ไม่เพิ่มมากเมื่อมีจำนวนผู้ประเมินหรือจำนวนแผนผังมโนทัศน์เพิ่มมากขึ้น) เมื่อกำหนดค่าสัมประสิทธิ์สรุปอ้างอิงเพื่อการนำไปใช้อิงเกณฑ์เท่ากับ 0.8 พบว่าวิธีการให้คะแนนของโนแวคและโกวินต้องใช้ผู้ประเมินจำนวน 2 คน และจำนวนแผนผังมโนทัศน์จำนวน 8 แผนผัง หรือผู้ประเมิน 1 คน แผนผังมโนทัศน์ 12 แผนผัง และวิธีการให้คะแนนของบอลท์ ต้องใช้ผู้ประเมิน 4 คน และจำนวนแผนผังมโนทัศน์ 12 แผนผัง หรือผู้ประเมิน 3 คน แผนผังมโนทัศน์ 16 แผนผัง ทั้งนี้ขึ้นอยู่กับผู้ปฏิบัติว่าจะเลือกใช้ให้เหมาะสมกับการเรียนการสอน

เพื่อแสดงให้เห็นการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการให้คะแนนแผนผังมโนทัศน์ที่ใช้วิธีการให้คะแนนทั้งสองรูปแบบตามรายละเอียดที่กล่าวมาข้างต้นอย่างชัดเจน ผู้วิจัยนำเสนอการเปรียบเทียบเป็นกราฟเส้น ดังแสดงในภาพประกอบ



แผนภาพที่ 4.1 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์จากการใช้วิธีการให้คะแนนของโนแวคและโกวิน ที่จำนวนแผนผังมโนทัศน์ 4-28 แผนผัง จำนวนผู้ประเมิน 1-6 คน ตามแบบแผนการวัด $P \times I \times R$



แผนภาพที่ 4.2 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์จากการใช้วิธีการให้คะแนนของบอลท์ ที่จำนวนแผนผังมโนทัศน์ 4-28 แผนผัง จำนวนผู้ประเมิน 1-6 คน ตามแบบแผนการวัด $P \times I \times R$

จากตารางที่ 4.21 และแผนภาพที่ 4.1- 4.2 พบว่าเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน ช่วงการตัดสินใจที่จำนวนแผนผังมโนทัศน์ 4-28 แผนผัง จำนวนผู้ประเมินตั้งแต่ 1-6 คน ให้ค่าสัมประสิทธิ์การสรุปอ้างอิงสูงกว่า วิธีการให้คะแนนของบอลท์ทุกช่วงค่า โดยวิธีการให้คะแนนของโนแวคและโกวินมีค่าระหว่าง 0.5828 ถึง 0.9825 และวิธีการให้คะแนนของบอลท์ มีค่าระหว่าง 0.2612 ถึง 0.9337 จากกราฟจะเห็นว่าเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน ความชันของกราฟเริ่มเปลี่ยนแปลงน้อยลง เมื่อจำนวนแผนผังมโนทัศน์ตั้งแต่ 12 แผนผังขึ้นไป สำหรับวิธีการให้คะแนนของบอลท์ ความชันของกราฟเริ่มเปลี่ยนแปลงน้อยลง เมื่อจำนวนแผนผังมโนทัศน์ตั้งแต่ 8 แผนผังขึ้นไป นั่นแสดงว่า ถึงแม้ว่าจำนวนชิ้นงานจะมากขึ้นก็ไม่มีผลต่อการเพิ่มค่าสัมประสิทธิ์สรุปอ้างอิงมากนัก

บทที่ 5

บทสรุป อภิปรายและข้อเสนอแนะ

การวิจัยในเรื่องนี้ ต้องการศึกษาคุณภาพของแผนผังมโนทัศน์ที่ใช้ในการประเมิน เมื่อใช้วิธีตรวจให้คะแนนที่แตกต่างกัน โดยทำการเปรียบเทียบระหว่างวิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ ซึ่งมีวัตถุประสงค์การวิจัย คือ

1. เพื่อศึกษาและพัฒนาเกณฑ์การให้คะแนนการประเมินแผนผังมโนทัศน์ด้วยวิธีของโนแวคและโกวิน และวิธีของบอลท์

2. เพื่อเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ เมื่อตรวจโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และใช้วิธีการให้คะแนนของบอลท์ โดยพิจารณาจาก

2.1 ความตรงตามสภาพ

2.2 ความตรงเชิงโครงสร้าง

2.3 ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ

2.4 ความเที่ยงซึ่งแทนด้วยสัมประสิทธิ์สรุปอ้างอิง

3. เพื่อวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสมเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน หรือวิธีการให้คะแนนของบอลท์

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นผู้ประเมินซึ่งเป็นครูสอนวิชาคณิตศาสตร์ชั้นมัธยมศึกษาปีที่ 1-3 และนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ ในการวิจัยครั้งนี้มีกลุ่มตัวอย่างที่ศึกษาประกอบด้วยกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2548 โรงเรียนสุรนารีวิทยา สำนักงานเขตพื้นที่การศึกษานครราชสีมา เขต1 สังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน กระทรวงศึกษาธิการ จำนวน 220 คน กลุ่มตัวอย่างผู้ประเมิน เนื่องจากในการวิจัยครั้งนี้ต้องใช้ครูผู้สอนวิชาคณิตศาสตร์ชั้น มัธยมศึกษาปีที่ 1-3 เป็นผู้ประเมินแผนผังมโนทัศน์ซึ่งต้องใช้เวลาและความสนใจ ตั้งใจในการประเมิน ผู้วิจัยทำการเลือกครูผู้ประเมินโดยวิธีการเลือกแบบเจาะจง (purposive sampling) เพราะต้องหาผู้ประเมินที่เต็มใจและมีเวลาช่วยประเมินอย่างเต็มที่ มีประสบการณ์ในการสอนอย่างน้อย 3 ปี ครูที่ได้รับเลือกเป็นกลุ่มตัวอย่าง รวมทั้งหมด 4 คน แบ่งครู 4 คน เป็น 2 กลุ่ม กลุ่มละ 2 คน โดยวิธีสุ่มอย่างง่ายโดยการจับสลาก ซึ่งผู้ประเมินกลุ่มที่ 1 จะใช้วิธีการให้คะแนนของโนแวคและโกวิน โดยผู้ประเมิน

ทุกคนตรวจให้คะแนนแผนผังที่ 1 และแผนผังที่ 2 และผู้ประเมินกลุ่มที่ 2 จะใช้วิธีการให้คะแนนของบอลท์ โดยผู้ประเมินทุกคนตรวจให้คะแนนแผนผังที่ 1 และแผนผังที่ 2 ดังนั้นผู้ประเมินแต่ละคนจะตรวจผลงานของนักเรียนจำนวน 220 คน จำนวน 440 ชิ้น

การวิเคราะห์ข้อมูล โดยวิเคราะห์ความตรงตามสภาพ จากค่าสัมประสิทธิ์สหสัมพันธ์ (r) และแปลงค่าความตรงตามสภาพจากค่าสัมประสิทธิ์สหสัมพันธ์ เป็นคะแนนฟิชเชอร์ซี (Z_r)โดยใช้สูตร (Glass and Stanley , 1970) เปรียบเทียบความเป็นอิสระของความตรงตามสภาพในภาพรวมโดยใช้สถิติทดสอบไคว์สแควร์ (χ^2) วิเคราะห์ความตรงเชิงโครงสร้าง (construct validity) ด้วยเมทริกซ์พหุลักษณะพหุวิธี(multitrait-multimethod : MTMM) วิเคราะห์ค่าฟังก์ชันสารสนเทศของข้อสอบ (item information function) และค่าฟังก์ชันสารสนเทศของแบบสอบ (test information function) ตามทฤษฎีการตอบสนองข้อสอบ วิเคราะห์คุณภาพด้านความเที่ยงโดยการวิเคราะห์ค่าสัมประสิทธิ์สรุปอ้างอิง(Generalizability Coefficient) โดยใช้โปรแกรม GENOVA 3.1 (GENeralized analysis Of VAriance System) ออกแบบการวัดเป็น two-facet design โดยมีรูปแบบเป็น $P \times I \times R$ crossed design กล่าวคือ (P) แทนจำนวนนักเรียน (I) แทนจำนวนแผนผังมโนทัศน์ และ (R) แทนจำนวนผู้ประเมิน ซึ่งการออกแบบการวัดครั้งนี้ มีผู้ประเมิน 2 คน ตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำของนักเรียนทุกคน โดยใช้วิธีการตรวจให้คะแนนแบบของโนแวคและโกวิน และผู้ประเมิน 2 คน ตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำของนักเรียนทุกคน โดยใช้วิธีการตรวจให้คะแนนแบบของบอลท์ และวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิงเพื่อการตัดสินใจเชิงสรุปอ้างอิง (D-STUDY) เลือกจำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสม

สรุปผลการวิจัย

ผลการวิจัยที่ได้แบ่งเป็น 2 ตอน คือ ตอนที่ 1 สรุปผลการวิเคราะห์ข้อมูลเบื้องต้น และตอนที่ 2 คือ สรุปผลการวิเคราะห์ข้อมูลเพื่อตอบคำถามการวิจัยตามวัตถุประสงค์

ตอนที่ 1 สรุปผลการวิเคราะห์ข้อมูลเบื้องต้น

ผลการวิเคราะห์ข้อมูลเบื้องต้นพบว่า ค่าร้อยละของค่าเฉลี่ย (Mean %) ของการให้คะแนนด้วยวิธีของโนแวคและโกวินมีค่าเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์ ในขณะที่สัมประสิทธิ์ของการกระจาย (C.V.) การให้คะแนนด้วยวิธีของบอลท์ มีค่าสูงกว่าวิธีการให้คะแนนของโนแวคและโกวิน และเมื่อเปรียบเทียบความแตกต่างของค่าเฉลี่ยเป็นรายคู่พบว่าวิธีการให้คะแนนของ

โนแวคและโกวินมีคะแนนเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

ตอนที่ 2 สรุปผลการวิเคราะห์ข้อมูลเพื่อตอบคำถามการวิจัยตามวัตถุประสงค์

2.1 ผลการศึกษาและพัฒนาวิธีการให้คะแนนการประเมินแผนผังมโนทัศน์ของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์

วิธีการให้คะแนนการประเมินแผนผังมโนทัศน์ได้พัฒนาแนวคิดมาจากงานวิจัยของโนแวคและโกวิน (Novak & Gowin, 1984) และงานวิจัยของบอลท์ (Bolte, 1999) ซึ่งมีรายละเอียดดังนี้

เกณฑ์การให้คะแนนของโนแวคและโกวิน จะพิจารณาแผนผังมโนทัศน์ในด้านคุณภาพ 4 ด้าน คือ พิจารณาจากประพจน์ (propositions) การจัดลำดับขั้นตอน (hierarchy) การเชื่อมระหว่างมโนทัศน์ (cross link) และตัวอย่าง (examples) ซึ่งแต่ละด้านจะมีคะแนนเต็ม 4 คะแนน

เกณฑ์การให้คะแนนของบอลท์ จะพิจารณาให้คะแนนในด้านการจัดแผนผังมโนทัศน์ (organization) 6 คะแนน ความถูกต้อง (accuracy) 4 คะแนน

2.2 ผลการเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ เมื่อตรวจโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และใช้วิธีการให้คะแนนของบอลท์

2.2.1 ค่าความตรงตามสภาพของคะแนนการประเมินแผนผังมโนทัศน์แบบสร้างผังจากคำ เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินมีค่าสูงกว่าการตรวจให้คะแนนด้วยวิธีของบอลท์

2.22 แผนผังมโนทัศน์แบบสร้างผังจากคำ เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน และการตรวจให้คะแนนด้วยวิธีของบอลท์ มีความตรงเชิงโครงสร้างทั้งสองวิธี

2.2.3 ค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ ของวิธีการให้คะแนนของโนแวคและโกวินมีค่าฟังก์ชันสารสนเทศของข้อสอบโดยเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์ วิธีการให้คะแนนของโนแวคและโกวินมีประสิทธิภาพสัมพัทธ์สูงกว่าวิธีการให้คะแนนบอลท์

2.2.4 ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินแผนผังมโนทัศน์แบบสร้างผังจากคำ เมื่อตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินมีค่าสูงกว่าการตรวจให้คะแนนด้วยวิธีของบอลท์

2.3 ผลการวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสม

เมื่อกำหนดผู้ประเมินจำนวน 1-6 คน จำนวนแผนผังมโนทัศน์ 4-28 แผนผัง การตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน มีค่าสัมประสิทธิ์การสรุปอ้างอิงอยู่ระหว่าง 0.5828 ถึง 0.9825 และเมื่อตรวจให้คะแนนด้วยวิธีของบอลท์ มีค่าสัมประสิทธิ์อ้างอิงอยู่ระหว่าง 0.2612 ถึง 0.9337 และพบว่าจำนวนผู้ประเมินและจำนวนแผนผังที่เหมาะสม ที่ทำให้ค่าสัมประสิทธิ์สรุปอ้างอิงอยู่ในเกณฑ์เหมาะสม (ไม่เพิ่มมากเมื่อมีจำนวนผู้ประเมินหรือจำนวนแผนผังมโนทัศน์เพิ่มมากขึ้น) เมื่อกำหนดค่าสัมประสิทธิ์สรุปอ้างอิงเพื่อการนำไปใช้อิงเกณฑ์ เท่ากับ 0.8 พบว่าวิธีการให้คะแนนของโนแวคและโกวินต้องใช้ผู้ประเมินจำนวน 2 คน และจำนวนแผนผังมโนทัศน์จำนวน 8 แผนผัง หรือผู้ประเมิน 1 คน แผนผังมโนทัศน์ 12 แผนผัง และวิธีการให้คะแนนของบอลท์ ต้องใช้ผู้ประเมิน 4 คน และจำนวนแผนผังมโนทัศน์ 12 แผนผัง หรือผู้ประเมิน 3 คน แผนผังมโนทัศน์ 16 แผนผัง ทั้งนี้ขึ้นอยู่กับผู้ปฏิบัติว่าจะเลือกใช้ให้เหมาะสมกับการเรียนการสอน

อภิปรายผลการวิจัย

จากผลการวิจัยที่กล่าวมาข้างต้นสามารถอภิปรายได้ใน 3 ประเด็น ได้แก่ ประเด็นแรก การศึกษาและพัฒนาวิธีการให้คะแนนการประเมินแผนผังมโนทัศน์ของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ ประเด็นที่สอง การเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ เมื่อตรวจโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และใช้วิธีการให้คะแนนของบอลท์ และประเด็นที่สาม การวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสมเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์

ประเด็นที่ 1 การศึกษาและพัฒนาวิธีการให้คะแนนการประเมินแผนผังมโนทัศน์ของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์

วิธีการให้คะแนนการประเมินแผนผังมโนทัศน์ได้พัฒนาแนวคิดมาจากงานวิจัยของโนแวคและโกวิน (Novak & Gowin, 1984) และงานวิจัยของบอลท์ (Bolte, 1999) พบว่าวิธีการให้คะแนนทั้ง 2 วิธีมีข้อดีและข้อเสียแตกต่างกัน โดยวิธีการให้คะแนนของโนแวคและโกวิน จะมีลักษณะการให้คะแนนแบบแยกองค์ประกอบ เพื่อให้เห็นคุณภาพของการปฏิบัติงานรายละเอียดปลีกย่อย การประเมินจึงมีการวิเคราะห์แยกมิติหรือองค์ประกอบที่จะพิจารณาให้ระดับคะแนน ซึ่งแต่ละมิติหรือองค์ประกอบก็จะแบ่งระดับตามระดับคุณภาพความสามารถ วิธีการให้คะแนนของโนแวคและโกวินจะต้องรวมคะแนนเป็นคะแนนรวมทั้งหมดของผลงาน การตรวจให้คะแนน

ด้วยวิธีนี้จึงต้องใช้เวลามาก แต่รายละเอียดของแต่ละระดับจะมีประโยชน์ เมื่อต้องการที่จะวินิจฉัยข้อผิดพลาดของผู้เรียน (Kinchin et al, 2000; Klein, Chung, Osmundson, Herl, & O'Neil, 2002; McClure, Sonak, & Suen, 1999; XiuFeng and Hinchey, 1996) สำหรับวิธีการให้คะแนนของบอลท์ มีลักษณะการให้คะแนนแบบภาพรวม ซึ่งวิธีการตรวจให้คะแนนของบอลท์จะง่ายและประหยัดเวลา แต่จะไม่ได้รายละเอียดข้อผิดพลาดที่ชัดเจน (Ruiz-Primo, Schultz, Li, & Shavelson, 2001; Yin et al, 2004; รุ่งนภา ทศภานนท์, 2544; วิยะดา ระวังสุข, 2545) การนำวิธีการตรวจให้คะแนนแต่ละวิธีไปใช้จำเป็นต้องใช้เวลาในการศึกษาให้เข้าใจ เพราะทักษะและความชำนาญในการตรวจให้คะแนนจะช่วยเพิ่มคุณภาพในการประเมินแผนผังมโนทัศน์ และสามารถนำสารสนเทศหรือข้อมูลย้อนกลับมาพัฒนาผู้เรียนได้อย่างเหมาะสมและมีประสิทธิภาพ เมื่อพิจารณาถึงค่าเฉลี่ยของคะแนนที่ได้จากการตรวจให้คะแนนของผู้ประเมินของวิธีการตรวจแต่ละวิธีพบว่า ผู้ประเมินที่ตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินให้ค่าเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์ เพราะว่า การพิจารณาแยกทีละองค์ประกอบในการตรวจให้คะแนนทำให้ได้รายละเอียดมากกว่า ทั้งในด้านข้อผิดพลาดหรือข้อมูลที่เป็นประโยชน์ ซึ่งการพิจารณาในภาพรวมทำให้ขาดข้อมูลที่เป็นประโยชน์ จึงส่งผลให้เมื่อตรวจให้คะแนนด้วยวิธีของบอลท์มีคะแนนเฉลี่ยต่ำและมีความแตกต่างจากวิธีการตรวจให้คะแนนโนแวคและโกวิน

ประเด็นที่ 2 การเปรียบเทียบคุณภาพของการประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ เมื่อตรวจโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และใช้วิธีการให้คะแนนของบอลท์

การคำนวณความตรงตามสภาพของแบบสอบที่มีวิธีการตรวจให้คะแนนที่แตกต่างกัน ซึ่งพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์ (r) ระหว่างคะแนนของผู้สอบที่ได้จากการสอบด้วยแบบสอบประเภทเลือกตอบประกอบด้วย 2 เรื่อง คือ สมบัติของจำนวนนับ และระบบจำนวนเต็ม ที่ผู้วิจัยสร้างขึ้น กับคะแนนที่ได้จากการประเมินแผนผังมโนทัศน์ โดยที่การตรวจให้คะแนนจากการประเมินแผนผังมโนทัศน์มีการตรวจให้คะแนน 2 วิธีที่แตกต่าง พบว่าค่าความตรงตามสภาพขึ้นอยู่กับวิธีการตรวจให้คะแนนในแต่ละวิธี โดยที่วิธีการตรวจให้คะแนนของโนแวคและโกวินมีค่าความตรงตามสภาพสูงกว่าวิธีการตรวจให้คะแนนของบอลท์ แสดงว่าการให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำด้วยวิธีทั้งสอง สามารถวัดความสามารถในการคิดของนักเรียนได้ตรงกับความสามารถที่แท้จริงของนักเรียนได้แตกต่างกัน ทั้งนี้อาจเป็นเพราะผู้ประเมินได้ศึกษาวิธีการตรวจให้คะแนนให้ชัดเจนก่อนนำไปตรวจให้คะแนนแผนผังมโนทัศน์ ซึ่งวิธีการตรวจให้คะแนนของโนแวคและโกวินจะมีความชัดเจนในรายละเอียดมากว่าวิธีการให้คะแนนของบอลท์ มีการ

วิเคราะห์แยกมิติหรือองค์ประกอบที่จะพิจารณาเป็นด้าน และความตรงตามสภาพที่ค่อนข้างสูงอาจเป็นเพราะรูปแบบการตอบ (response format) ของแผนผังมโนทัศน์ โดยรูปแบบที่นำมาใช้ในการศึกษาครั้งนี้เป็นแผนผังมโนทัศน์แบบสร้างผังจากคำ (construct a map) ซึ่งหยินและคณะ (Yin et al, 2004) พบว่าแผนผังมโนทัศน์รูปแบบนี้สามารถนำไปใช้สะท้อนความคิดรวบยอด และความเข้าใจเนื้อหาในบทเรียนของนักเรียนได้ดีกว่าแบบอื่น จึงส่งผลให้ค่าความตรงตามสภาพของวิธีการตรวจให้คะแนนของทั้งสองวิธีมีค่าสูง

ผลการตรวจสอบความตรงของการตรวจให้คะแนนแผนผังมโนทัศน์ด้วยวิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์ ในส่วนนี้จะเป็นการตรวจสอบความตรงเชิงโครงสร้างโดยใช้เมทริกซ์พหุลักษณะพหุวิธี (Multitrait-Multimethod Matrix) ประกอบด้วยการวัดคุณลักษณะ 2 อย่าง (สมบัติของจำนวนนับ และ ระบบจำนวนเต็ม) ด้วยวิธีการวัด 3 วิธี (แบบทดสอบประเภทเลือกตอบ, แผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของโนแวคและโกวิน และแผนผังมโนทัศน์ที่มีการตรวจให้คะแนนโดยใช้วิธีการให้คะแนนของบอลท์) พบว่า ทั้งสองวิธีมีความตรงเชิงโครงสร้าง เนื่องจากวิธีการวัดหลายวิธีจะลู่เข้า (converge) สำหรับการวัดลักษณะเดียวกันกล่าวคือ คะแนนจากการวัดลักษณะเดียวกัน ถึงแม้จะใช้วิธีการวัดต่างกันย่อมมีความสัมพันธ์กันสูง และวิธีการวัดจะต้องจำแนก (discriminate) สำหรับการวัดลักษณะต่างกัน กล่าวคือ คะแนนจากการวัดคุณลักษณะต่างกันไม่ว่าจะใช้วิธีการวัดเดียวกันหรือต่างกัน ย่อมมีความสัมพันธ์กันต่ำ ซึ่งแสดงถึงความตรงเชิงโครงสร้างของการตรวจให้คะแนนในแต่ละวิธีที่นำมาศึกษาในครั้งนี้ ซึ่งผลการวิเคราะห์คุณภาพความตรงเชิงโครงสร้างพบว่าวิธีตรวจการให้คะแนนของโนแวคและโกวินมีความตรงเชิงโครงสร้างสูงกว่าวิธีการตรวจให้คะแนนของบอลท์ ซึ่งความสามารถในการจำแนกผู้ที่รู้หรือไม่รู้ และการตรวจให้คะแนนนั้นแสดงถึงระดับความรู้ที่แตกต่างกัน จึงทำให้ความตรงของคะแนนที่ได้แตกต่างกัน ซึ่งสอดคล้องกับงานวิจัยของเคน และคณะ (Klein et al, 2002) ได้ทำการทดสอบความตรงของแผนผังมโนทัศน์ในการใช้เป็นเครื่องมือสำหรับการประเมินในวิชาวิทยาศาสตร์ โดยใช้เมริกซ์พหุลักษณะพหุวิธี ประกอบด้วยการวัดคุณลักษณะ 2 อย่าง (การมองเห็น และ การได้ยิน) ด้วยวิธีการวัด 3 วิธี (แผนผังมโนทัศน์แบบสร้างผังจากคำ, แบบทดสอบประเภทความเรียง, แบบทดสอบประเภทเลือกตอบ) ที่พบว่าวิธีการวัดด้วยแผนผังมโนทัศน์แบบสร้างผังจากคำ มีความไวต่อระดับความสามารถของนักเรียน จึงมีความตรงเชิงโครงสร้างแตกต่างจากวิธีการวัดด้วยแบบทดสอบประเภทความเรียงและแบบทดสอบประเภทเลือกตอบ

การเปรียบเทียบค่าฟังก์ชันสารสนเทศของข้อสอบและแบบสอบ ระหว่างวิธีที่ตรวจให้คะแนนที่แตกต่างกัน พบว่าวิธีการให้คะแนนของโนแวคและโกวินมีค่าฟังก์ชันสารสนเทศของข้อสอบโดยเฉลี่ยสูงกว่าวิธีการให้คะแนนของบอลท์ อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 และค่าฟังก์ชันสารสนเทศของแบบสอบสูงกว่าวิธีการให้คะแนนของบอลท์ ซึ่งธีการให้คะแนนของโนแวคและโกวินมีประสิทธิภาพสัมพัทธ์สูงกว่าวิธีการให้คะแนนของบอลท์ แสดงว่าการตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินทำให้ได้คะแนนที่ประมาณค่าความสามารถประมาณค่าความสามารถของผู้สอบได้ใกล้เคียงกับความสามารถที่แท้จริงได้มากกว่า ทั้งนี้อาจเป็นเพราะว่าวิธีการตรวจให้คะแนนของโนแวคและโกวินมีลำดับขั้นในการตอบมากกว่าวิธีการให้คะแนนของบอลท์ ซึ่งจำแนกระดับความสามารถของของผู้ตอบได้มากกว่า และลำดับขั้นของการตอบข้อสอบแต่ละข้อสอดคล้องกับระดับความสามารถของผู้ตอบ จึงทำให้ประมาณค่าความสามารถของผู้สอบได้แม่นยำกว่า จึงทำให้ค่าฟังก์ชันสารสนเทศของแบบสอบสูงกว่าวิธีการให้คะแนนของบอลท์ นอกจากนี้เมื่อพิจารณาลงรายละเอียดของค่าฟังก์ชันสารสนเทศในระดับความสามารถของนักเรียนที่แตกต่างกัน พบว่าวิธีการให้คะแนนของโนแวคและโกวินให้ค่าฟังฟันก์สารสนเทศสูงในระดับความสามารถปานกลางถึงค่อนข้างอ่อน (-1 ถึง 0) ในขณะที่วิธีการตรวจให้คะแนนของบอลท์ให้ค่าฟังชันสารสนเทศสูงในระดับความสามารถค่อนข้างสูง (1 ถึง 2) แสดงให้เห็นว่าการตรวจให้คะแนนที่พิจารณารายละเอียดแบ่งย่อยเป็นมิติจะช่วยให้ผู้เรียนได้คะแนนมากกว่าการตรวจแบบภาพรวม ดังนั้นวิธีการตรวจให้คะแนนของโนแวคและโกวินจึงเหมาะที่จะใช้กับกลุ่มนักเรียนที่มีความสามารถปานกลางค่อนข้างต่ำ และวิธีการตรวจให้คะแนนของบอลท์เหมาะที่จะใช้กับกลุ่มนักเรียนที่มีความสามารถค่อนข้างสูง

คะแนนจากการตรวจให้คะแนนแผนผังมโนทัศน์ด้วยวิธีที่แตกต่างกัน 2 วิธี ผู้วิจัยได้ทำการประมาณค่าสัมประสิทธ์สรุปอ้างอิงของคะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำโดยศึกษาองค์ประกอบ 3 องค์ประกอบ ได้แก่ นักเรียน (P) แผนผังมโนทัศน์ (I) และผู้ประเมิน (R) $P \times I \times R$ ผู้ประเมิน 2 คน ตรวจให้คะแนนแผนผังมโนทัศน์ของนักเรียนทุกคน พบว่าค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์ของการตรวจให้คะแนนด้วยวิธีของโนแวคและโกวินมีค่าสูงกว่าการตรวจให้คะแนนด้วยวิธีของบอลท์ เมื่อพิจารณาวิธีการตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน พบว่าแหล่งความแปรปรวนที่ส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงที่ชัดเจนที่สุดคือ ความแปรปรวนของนักเรียน ซึ่งแสดงว่าความสามารถ ความรู้และความเข้าใจในการสร้างแผนผังมโนทัศน์ของนักเรียนแต่ละคนที่แตกต่างกันจะส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง ซึ่งสอดคล้องกับงานวิจัยของ McClure et al (1999) ที่พบว่าประสิทธิภาพหรือศักยภาพในการสร้าง

แผนผังมโนทัศน์ของนักเรียนแต่ละคนมีผลต่อความเที่ยงที่ได้จากการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง เพราะการที่นักเรียนมีสามารถในการสร้างแผนผังมโนทัศน์ต่างกันทำให้เกิดความคลาดเคลื่อนในการวัด นอกจากนี้ยังสอดคล้องกับงานวิจัยของรุซ-พรีโม และคณะ (Ruiz-Primo et al, 2001) ที่พบว่าแหล่งความแปรปรวนหลักที่มีค่าต่อสัมประสิทธิ์ความเที่ยงของแผนผังมโนทัศน์ก็คือความสามารถในการสร้างแผนผังของนักเรียน นอกจากนี้ยังพบว่าผู้ประเมินก็เป็นแหล่งความแปรปรวนที่ส่งผลต่อค่าสัมประสิทธิ์สรุปอ้างอิงของคะแนนแผนผังมโนทัศน์เช่นกัน ทั้งนี้อาจเป็นเพราะความชัดเจนของวิธีการตรวจให้คะแนน และวิธีการตรวจทำหลายขั้นตอนทำให้ผู้ประเมินตัดสินต่างกัน ซึ่งเกณฑ์การให้คะแนนของบอลท์นั้นยังมีความไม่ชัดเจน ผู้ประเมินจึงตรวจให้คะแนนได้แตกต่างกันไปตามดุลพิจนิจของตนเอง และมีข้อสังเกตอีกประการหนึ่งที่พบว่าน่าจะส่งผลต่อการตรวจให้คะแนนนั่นคือ ลายมือและการลากเส้นเชื่อมโยงระหว่างคำมโนทัศน์ที่ไม่ชัดเจนของนักเรียนส่งผลให้ผู้ประเมินตัดสินได้ต่างกัน

และเมื่อพิจารณาในภาพรวมแล้ว พบว่าวิธีการตรวจให้คะแนนของโนแวคและโกวินให้ค่าความตรงและความเที่ยงสูงกว่าวิธีการให้คะแนนของบอลท์ ทั้งนี้อาจเป็นเพราะ ความชัดเจนของวิธีการตรวจให้คะแนนของทั้งสองวิธี และวิธีการตรวจทำหลายขั้นตอนทำให้ผู้ประเมินตัดสินต่างกัน ซึ่งเกณฑ์การให้คะแนนของบอลท์นั้นยังมีความไม่ชัดเจน ผู้ประเมินจึงตรวจให้คะแนนได้แตกต่างกันไปตามดุลพิจนิจของตนเอง ในขณะที่วิธีการให้คะแนนของโนแวคและโกวินมีเกณฑ์การพิจารณาที่ละเอียด ชัดเจนกว่า จึงทำให้ผู้ประเมินจึงตรวจให้คะแนนได้สอดคล้องกันมากกว่าส่งผลให้คุณภาพด้านความตรงและความเที่ยงที่ได้จากการตรวจให้คะแนนแผนผังมโนทัศน์สูงกว่าวิธีการให้คะแนนของบอลท์

ประเด็นที่ 3 การวิเคราะห์จำนวนผู้ประเมินและจำนวนแผนผังมโนทัศน์ที่เหมาะสมเมื่อใช้วิธีการให้คะแนนของโนแวคและโกวิน และวิธีการให้คะแนนของบอลท์

การศึกษาเพื่อการตัดสินใจ (D-STUDY) จำนวนแผนผังมโนทัศน์ และจำนวนผู้ประเมินที่เหมาะสม ขึ้นอยู่กับการตัดสินใจของผู้ปฏิบัติที่สามารถเลือกช่วงการตัดสินใจใช้วิธีการให้คะแนนวิธีใด ที่จะนำไปปฏิบัติจริงแล้วได้ค่าความเที่ยงตามต้องการ เช่น ถ้าผู้สอน 4 คน ต้องการค่าความเที่ยง 0.8 สำหรับวิธีการตรวจให้คะแนนของโนแวคและโกวิน ต้องใช้แผนผังมโนทัศน์อย่างน้อย 12 แผนผัง จะเห็นว่ามีจำนวนแผนผังมโนทัศน์ที่สูง ในความเป็นจริงอาจกระทำได้ แต่ไม่สะดวกต้องใช้เวลานาน ดังนั้นผู้สอนอาจปรับค่าความเที่ยงเป็น 0.6 แทน ซึ่งต้องใช้แผนผังอย่างน้อย 8 แผนผัง ทำให้สะดวกและประหยัดเวลาได้ โดยสรุปแล้ว การพิจารณาว่าจะ

เพิ่มจำนวนแผนผังมโนทัศน์หรือจำนวนผู้ประเมินนั้น จะต้องพิจารณาว่าสามารถปฏิบัติได้จริง มีความสะดวกและประหยัด ในการเรียนการสอนเราสามารถปฏิบัติได้จริง ซึ่งในชั้นเรียนนั้นเราสามารถเพิ่มจำนวนแผนผังได้มากกว่าจำนวนผู้ประเมิน โดยเฉพาะการใช้ครูจำนวนมาก เป็นสิ่งที่ทำได้ยากสำหรับการประเมินในโรงเรียน แต่การเพิ่มชิ้นงานสามารถทำได้ง่ายกว่า ซึ่งสอดคล้องกับงานวิจัยของ (Yin et al, 2004) ที่พบว่าการเพิ่มจำนวนแผนผังมโนทัศน์หรือจำนวนผู้ประเมินนั้นต้องพิจารณาในเรื่องของเวลา ค่าใช่จ่าย ความสะดวกในการเรียนการสอน

ข้อเสนอแนะจากการวิจัยครั้งนี้

1. ควรสร้างความเข้าใจกับผู้ประเมินเกี่ยวกับเกณฑ์การตรวจให้คะแนนของแต่ละวิธี โดยที่ครูผู้ประเมินมีความรู้ในเรื่องที่ตรวจอย่างดี และชี้แจงถึงวิธีและเกณฑ์การตรวจให้เข้าใจตรงกัน เพื่อที่จะสามารถประเมินแผนผังมโนทัศน์ได้สอดคล้องกัน

2. ในการตรวจให้คะแนนแผนผังมโนทัศน์ ควรมีแผนผังมโนทัศน์เกณฑ์หรือเฉลยที่สมบูรณ์ไว้เป็นตัวเกณฑ์เทียบ

3. ควรมีการเตรียมความพร้อมให้กับนักเรียน โดยการอบรมให้ความรู้หรือนำแผนผังมโนทัศน์ไปใช้จัดกิจกรรมการเรียนการสอนในวิชาคณิตศาสตร์ โดยให้นักเรียนได้ฝึกสร้างแผนผังมโนทัศน์จนเกิดความชำนาญและเข้าใจวิธีการสร้างอย่างถ่องแท้ก่อนทำการประเมิน เพื่อนักเรียนจะสามารถถ่ายทอดความคิดออกมาได้อย่างมีลำดับขั้นตอน

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรศึกษาวิธีการให้คะแนนของบอลท์ เพื่อตรวจสอบยืนยันข้อมูลที่ได้จากงานวิจัยครั้งนี้

2. ควรศึกษาการประเมินแผนผังมโนทัศน์โดยใช้วิธีการให้คะแนนแบบอื่นๆ หรือให้ผู้เรียนเป็นผู้สร้างวิธีการให้คะแนน และประเมินผลงานนั้นด้วยวิธีที่สร้างขึ้น

3. ควรศึกษาองค์ประกอบอื่นๆที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง เช่น จำนวนครั้งในการตรวจให้คะแนน เพราะจะได้ตรวจสอบ ยืนยันข้อมูลที่ได้จากงานวิจัยครั้งนี้

3. ควรศึกษาแหล่งความคลาดเคลื่อนอื่นนอกจากการตรวจให้คะแนน เช่นรูปแบบการตอบ ที่ส่งผลต่อคุณภาพของการประเมินแผนผังมโนทัศน์ทั้งในด้านความตรงและความเที่ยง

4. ควรศึกษาความเป็นไปได้ของการตรวจให้คะแนนแผนผังมโนทัศน์แบบสร้างผังจากคำด้วยคอมพิวเตอร์หรือเครื่องตรวจให้คะแนนอื่นๆ

รายการอ้างอิง

ภาษาไทย

กนกวรรณ รัตนธน. (2544). *การเปรียบเทียบความตรงตามสภาพและความคงที่ในการประมาณค่าความสามารถของผู้สอบในการทดสอบแบบปรับเหมาะกับความสามารถของผู้สอบด้วยคอมพิวเตอร์เมื่อระดับความสามารถของผู้สอบและอัตราการใช้ข้อสอบซ้ำแตกต่างกัน*. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

กัญญา ทิพย์ลาย. (2545). *การสร้างแบบฝึกการเขียนสรุปความจากร้อยแก้ว*. วิทยานิพนธ์ปริญญามหาบัณฑิต ภาควิชาภาษาไทย บัณฑิตวิทยาลัย มหาวิทยาลัยทักษิณ.

กำพล ดำรงค์วงศ์. (2540). *การพัฒนาแบบจำลองคอมพิวเตอร์ช่วยสอนแบบอัจฉริยะเพื่อการสร้างแผนผังมโนทัศน์*. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ณัฐยากร แกล้วทนงค์. (2545). *ผลของการใช้เทคนิคแผนผังความคิดต่อผลสัมฤทธิ์ทางการอ่านของนักเรียนที่มีระดับความสามารถต่างกัน*. วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาจิตวิทยาการศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์.

ธนวัฒน์ แสนสุข. (2538). *การใช้จีอาร์เอ็ม จีพีซีเอ็ม และโมเดลโลจิสติก ในการเปรียบเทียบฟังก์ชันสารสนเทศของแบบวัดที่มีวิธีการให้คะแนนต่างกัน*. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ธัญพันธ์ พงศาวกุล. (2545). *ผลของการใช้แผนผังมโนทัศน์ที่มีต่อผลสัมฤทธิ์ทางการเขียนของนักเรียนชั้นประถมศึกษาปีที่ 6*. วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาการประถมศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์.

นงลักษณ์ วิรัชชัย. (2537). *ความสัมพันธ์โครงสร้างเชิงเส้น(LISREL) : สถิติวิเคราะห์สำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์*. กรุงเทพมหานคร : จุฬาลงกรณ์มหาวิทยาลัย.

บุญประกอบ จายค่าย. (2545). *การสร้างแบบวัดภาคปฏิบัติสำหรับโปรแกรมประมวลคำไมโครซอฟท์เวิร์ด*. วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี.

ประกิต ผลมูล. (2546). *การใช้แผนภูมิมโนมิติในการปรับเปลี่ยนมโนมติและมโนมติที่คลาดเคลื่อนในวิชาชีววิทยาเรื่องรากและลำต้น ของนักเรียนชั้นมัธยมศึกษาปีที่ 5*. วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาวิทยาศาสตร์ศึกษา บัณฑิตวิทยาลัย มหาวิทยาลัยสงขลานครินทร์.

ปวีณา ปีอาทิตย์. (2545). *การศึกษาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน.* วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ปิยวดี คงช่วย. (2544). *การพัฒนาแบบทดสอบความคิดวิจารณญาณ โดยใช้ทฤษฎีการตอบสนองข้อสอบสำหรับนักเรียนชั้นประถมศึกษาปีที่ 6.* วิทยานิพนธ์ปริญญามหาบัณฑิต มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี.

พรทิพย์ เพชรกาศ. (2543). *การเปรียบเทียบค่าฟังก์ชันสารสนเทศของแบบทดสอบเลือกตอบแบบถูกผิด ที่ตรวจให้คะแนนต่างวิธี.* วิทยานิพนธ์ปริญญามหาบัณฑิต มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี.

พรรณิทพย์ ม้ามณี. (2532). *การสอนคณิตศาสตร์แนวใหม่ระดับมัธยมศึกษา.* กรุงเทพมหานคร: สารศึกษาการพิมพ์.

มนัส บุญประกอบ. (2542). แผนภูมิมโนทัศน์กับการวิจัย ตอนที่ 1. *จดหมายข่าวสถาบันวิจัยพฤติกรรมศาสตร์* 17 (29 ตุลาคม 2542): 3-4.

เยาวดี วิบูลย์ศรี. (2539). *การวัดและการสร้างแบบสอบวัดผลสัมฤทธิ์.* กรุงเทพมหานคร : จุฬาลงกรณ์มหาวิทยาลัย.

รุ่งนภา ทศภานนท์. (2544). *ผลของการใช้แผนผังมโนทัศน์ที่มีต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์และความสามารถในการสร้างแผนผังมโนทัศน์ทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 4.* วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

วิยะดา ระวังสุข. (2545). *การประเมินความคิดรวบยอดวิชาวิทยาศาสตร์โดยใช้แผนผังมโนทัศน์.* วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

สมพร สุทัศนีย์. (2545). *การทดสอบทางจิตวิทยา.* กรุงเทพ; สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สาลิกา เมธนาวิน. (2545). *องค์ประกอบความฉลาดทางอารมณ์ของพยาบาลในการปฏิบัติการพยาบาล.* ปริญญาการศึกษาดุษฎีบัณฑิต สาขาวิจัยและพัฒนาหลักสูตร มหาวิทยาลัยศรินทรวิโรฒ ประสานมิตร.

สุชา จันทร์เอม. (2540). *จิตวิทยาทั่วไป.* กรุงเทพมหานคร: ไทยวัฒนาพาณิช.

สุทธิพร วงค์วิชัย. (2545). *ผลของการใช้เทคนิคกลุ่มร่วมมือโดยเน้นการสร้างแผนภูมิมโนทัศน์ที่มีต่อความคิดรวบยอดในการเรียนวรรณคดีไทยของนักเรียนชั้นประถมศึกษาปีที่ 5.* วิทยานิพนธ์ปริญญามหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยเชียงใหม่.

สุนีย์ สอนตระกูล. (2536). *การพัฒนาระบบการเรียนการสอนแบบจัดกรอบมโนทัศน์สำหรับวิชาชีววิทยาระดับชั้นมัธยมศึกษาตอนปลาย*. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต ภาควิชาหลักสูตรและการสอน บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

สุพิมพ์ ศรีพันวรสกุล. (2538). *การวิเคราะห์ความตรงลู่เข้าของการวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ตามทฤษฎีการอ้างอิงสรุป*. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

สุรพันธ์ ตันศรีวงษ์. (2538). *วิธีการสอน*. กรุงเทพมหานคร: สกายบุกส์.

ศิริชัย กาญจนวาสี. (2535). *เอกสารการสอนชุดวิชาการพัฒนาแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน*.นนทบุรี: โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

ศิริชัย กาญจนวาสี. (2538). *รวมบทความวิชาทฤษฎีการวัดและการประเมิน*. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2542). *โมเดลการตอบสนองข้อสอบแบบพหุภาค*. เอกสารประกอบการสอนวิชา 2702840 ทฤษฎีการวัดขั้นสูง มปท.

ศิริชัย กาญจนวาสี. (2544). *ทฤษฎีการทดสอบแบบดั้งเดิม*. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศิริชัย กาญจนวาสี. (2544). *ทฤษฎีการทดสอบแนวใหม่*. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

สุพัฒน์ สุกมลสันต์. (2539). *ธนาคารข้อทดสอบและการทดสอบปรับเปลี่ยนด้วยคอมพิวเตอร์*. กรุงเทพมหานคร: บริษัทพิมพ์ดี จำกัด.

อุทุมพร จามรมาน. (2537). *ทฤษฎีการวัดทางจิตวิทยา*. กรุงเทพมหานคร : ฟันนี่พับบลิชชิ่ง.

เอมอร จังศิริพรปกรณ์. (2545). *การเปรียบเทียบคุณภาพของแบบสอบเลือกตอบเมื่อตรวจด้วยวิธีการให้คะแนนความรู้บางส่วนกับวิธีประเพณีนิยม*. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Allen, M.J. and Yen, W.M. (1979). *Introduction to Measurement Theory.* California : Brooks/Cole Publishing Company.

Baroody, A.J.; & Bartels, B.H. (2001). Assessing understanding in mathematics with concept mapping. *Mathematics in School*. 30(3):24-27

Bolte, L.A. (1998). *Using concept maps and interpretive essays for assessment in mathematics.* Journal School Science & Mathematics in School. 30(3): 24-27.

Brennan, R.L. (1983). *Elements of Generalizability Theory*. Lowa : ACT Publications.

Campbell, D.T., and Fiske, D.W. (1959). Convengent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*. 56:81-105.

Carmines, E.G., and Zeller. R.A. (1987). *Reliability and Validity Assessment*. California : Sage Publications.

Cassidy, A.; Griffiths, T.;& Nakonechny, J. (2001). Concept mapping: Mirroring process of thinking and learning. [Online]. Available from: http://trochim.human.cornell.edu/research/ccp/tcands.htm.

Cooney, Thomas J.,Davis, Edward J. and Henderson, K.B. (1975). *Dynamics of Teaching Secondary School Mathematics*. Boston : Houghton Mifflin Company.

Cronbach, L.J., and Meehl, P.E. (1955). Construct Validity in Psychological Test. *Psychological Bulletin*. 52 :281-302.

Cronbach, L.J., et al. (1972). *The Dependability of Behavioral Measurements*. New York : John Wiley.

Donoghue, J. R. (1994). An Empirical Examination of the IRT Information of Polytomously Scored Reading Item Under The Generalized Partial Credit Model. *Journal of Educational Measurement,* 31. (4) : 295-311.

Donovan, Johnson A. and Derald, Rising R. (1972). *Guideline for Teaching Mathematics Secon Edition*. Wadsworth Publishing.

Eggen, Paul D. and Kauchak, Donald P. (1975). *Stratigies for Teaching Content and Thinking Skill*. Boston : Allyn and Bacon.

Goodwin, William L. and Klausmeier, Herbert J. (1995). *Falititating Student Learning : An Introduction to Educational Psychology*. New York : Harrer & Row.

Hambleton,R.K. and Swaminathan. H. (1985). *Item Response Theory:Principles and Application***.** Boston: Kluner-Nijhoff Publishing.

Hambleton, R. K., and Zaal, J. N. (1991). *Advanced in Educational and Psychological Testing: Theory and Applications.* Massachusetts: Kluwer Academic Publishers.

Hart, D. (1994). *Authentic Assessment: a hand book for educators.* New York: Addison-Wesley.

Kane, M.T. (1982). A Sampling Model for Validity. *Applied Psychological Measurement*. 6 : 125-160.

Kerlinger, F.N. (1986). *Foundation of Behavioral research*. 3rd.ed. Tokyo : CBS Publishing Japan.

Klein, D.C.D.,Chung, G.K.W.K., Osmundson, E., Herl, H.E., & O'Neil, H.F. (2002). *Exmining the validity of knowledge mapping as a measure of elementary students' scientific understanding* (CSE Tech Rep. No. 557). Los Angeles: University of California, National Center for Research on Evaluation, Standard, and Student Testing.

Koch, W. R. (1983). Likert Scaling Using the Graded Response Latent Trait Model. *Applied Psychological Measurement***.** 7 (1) : 15-32.

Lanzing, J.W.A. (2000). *The concept mapping homepage*. [Online]. Available from: http://usres.edte.utwente.nl/lanzing/cmhome.htm.

McCabe, D.(1995). The concept mapping workshop. [Online]. Available from: http://158.132.100.221/CMWkshp_golder/Cncpt_Mapp.Wkshop.html

McClure, J. R., Sonak, B., & Suen, H.K. (1999). Cencept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*. 36 : 475-492.

McCown, R. and Roup, Peter. (1992). *Educational Psychology and Classroom Practice : A Partnership*. Boston :Allyn an Bacon.

Muraki,E. and Bock,R.D. (1993). *Parscale.* Scientific Software,Inc.USA.

Novak, J.D. and Gowin, D. B. (1984). *Learning how to learn*. New York: Cambrigde University Press.

Novak, J.D. (2001). *The theory underlying concept maps and how to construct them.* [Online]. Available from: http://cmap.cognt.uwf.edu.info.

Nunnally, J.C. (1978). *Psychometric Theory*. New York : McGrow-Hill.

Plotnick, Eric. (1997). *Concept mapping: a graphical system for understanding the relationship between concepts*. [Online]. Available from: http://www.car.chula.ac.th.

Pressley, Michael and McCormick, Christine B. (1995). *Advance Education Psychology for Educator, Researcher and Policymakers*. New York : Harper Collings.

Reise, S. P., and Yu, J. (1990). Parameter Recovery in the Graded Response Model Using MULTILOG. *Journal of Educational Measurement.* 27 (2) : 133-144.

Ruiz-Primo, M.A.; & Shavelson, R.J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*. 33(6) : 569-600.

Ruiz-Primo, M. A., Schultz, S.E., Li, M., & Shavelson, R. J. (2001). *Comparison of the reliability and scores from two concept-mapping technique*. Journal of Research in Science Teaching. 38 : 260-278.

Schmitt, N., and Stults, D.M. (1986). Methodology Review : Analysis of Multitraui-Multimethod Matrices. *Applied Psychological Measurement*. 10 : 1-22.

Trochim, M.K. William. (1997). *An introduction to concept mapping for planning Evaluation*. [Online]. Available from: http://trochim.human.cornell.edu/research/epp1/epp1/htm

Yue Yin, Vandies, Ruiz-Primo, M. A., Alaya and Shavelson. (2004). *A comparison of two construct-a-concept map science assessment: created linking phrases and selected linking phrases*.(CSE Rep. No.624). Los Angeles: University of California, National Center for Research on Evaluation, Standard, and Student Testing (CRESST).

Zeilik, M. (2000). *Classroom assessment techniques concept mapping.* [Online]. Available from: http://www.flaguide.org/cat/minutepapers/conmap1.htm

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายชื่อผู้เชี่ยวชาญ

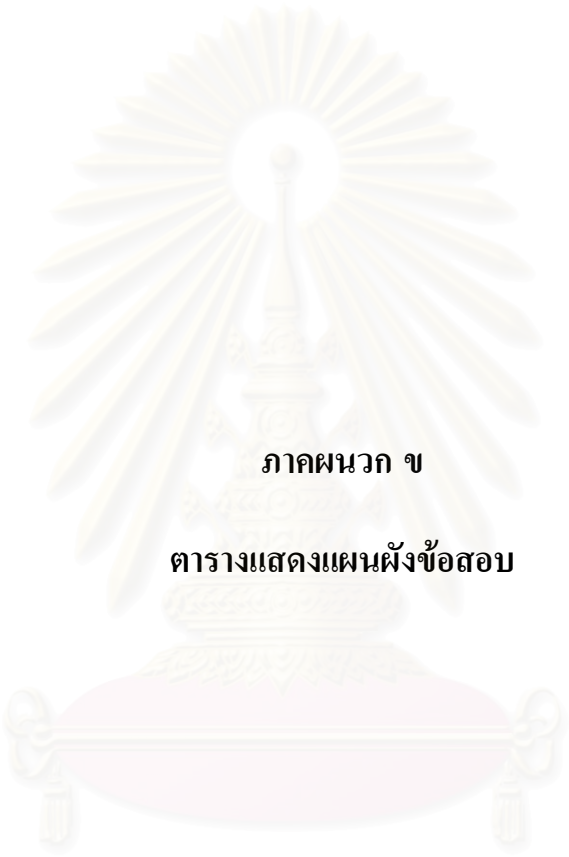
รายชื่อผู้เชี่ยวชาญทางการวัดและประเมินผลการศึกษา

1. รองศาสตราจารย์ ดร. ดวงกมล ไตรวิจิตรคุณ	อาจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. ผู้ช่วยศาสตราจารย์ ดร. ณัฏฐภรณ์ หลาวทอง	อาจารย์ประจำภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3. ผู้ช่วยศาสตราจารย์ โกวิท เชื่อมกลาง	อาจารย์ประจำภาควิชาทดสอบและวิจัยการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์

รายชื่อผู้เชี่ยวชาญทางด้านกลุ่มสาระการเรียนรู้คณิตศาสตร์

1. รองศาสตราจารย์ ดร. อัมพร ม้าคนอง	อาจารย์ประจำภาควิชาหลักสูตรการสอน และเทคโนโลยีการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2. อาจารย์ขนิษฐาน คำทอน	อาจารย์กลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
3. อาจารย์วิมลมาศ อำพลพงษ์	อาจารย์กลุ่มสาระการเรียนรู้คณิตศาสตร์ โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ข

ตารางแสดงแผนผังข้อสอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 1 แผนผังข้อสอบ (Testing Blueprint) หน่วยการเรียนรู้เรื่อง สมบัติของจำนวนนับ

ตารางแสดงแผนผังข้อสอบ (Testing Blueprint)

วิชา คณิตศาสตร์ ช่วงชั้นที่ 3 หน่วยการเรียนรู้เรื่อง สมบัติของจำนวนนับ

หลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 การสอบวัดด้านความสามารถ

ระดับชั้นมัธยมศึกษาปีที่ 1 จุดมุ่งหมายของการทดสอบ : เปรียบเทียบคุณภาพของการประเมิน

เวลาที่ใช้ในการทดสอบ 60 นาที ผู้วางแผน นางสาวทัตมณี ชูขวัญ

สาระการเรียนรู้	ผลการเรียนรู้ที่คาดหวัง	นำหนักรายข้อ	
		คิดเป็น%	จำนวนข้อ
1.จำนวนนับ	1. 1 เมื่อกำหนดจำนวนนับให้ นักเรียนสามารถหาตัวประกอบได้	15.38	2
	1.2 เมื่อกำหนดจำนวนนับให้ นักเรียนสามารถบอกได้ว่าจำนวนใดเป็นจำนวนคู่ และจำนวนใดเป็นจำนวนคี่	15.38	2
	1.3 เมื่อกำหนดจำนวนนับให้ นักเรียนสามารถบอกได้ว่าจำนวนใดบ้างเป็นจำนวนเฉพาะ จำนวนใดบ้างไม่เป็นจำนวนเฉพาะ	7.69	1
	1.4 นักเรียนสามารถหาตัวประกอบเฉพาะของจำนวนนับใดๆได้	7.69	1
2. ตัวหารร่วมมากและตัวคูณร่วมน้อยและการนำไปใช้	2.1 เมื่อกำหนดจำนวนนับตั้งแต่ 2 จำนวนขึ้นไป นักเรียนสามารถหา ห.ร.ม. ของจำนวนเหล่านั้นได้และนำความรู้เรื่อง ห.ร.ม. มาใช้แก้โจทย์ปัญหาได้	23.08	3
	2.2 เมื่อกำหนดจำนวนนับตั้งแต่ 2 จำนวนขึ้นไป นักเรียนสามารถหา ค.ร.น. ของจำนวนเหล่านั้นได้และนำความรู้เรื่อง ค.ร.น. มาใช้แก้โจทย์ปัญหาได้	15.38	2
	2.3 นำความรู้เรื่อง ห.ร.ม. และ ค.ร.น. ไปประยุกต์ใช้ได้	15.38	2
	รวม	100	13

ตารางภาคผนวกที่ 2 แผนผังข้อสอบ (Testing Blueprint) หน่วยการเรียนรู้เรื่อง ระบบจำนวนเต็ม

ตารางแสดงแผนผังข้อสอบ (Testing Blueprint)

วิชา คณิตศาสตร์ ช่วงชั้นที่ 3 หน่วยการเรียนรู้เรื่อง ระบบจำนวนเต็ม

หลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ. 2544 การสอบวัดด้านความสามารถ

ระดับชั้นมัธยมศึกษาปีที่ 1 จุดมุ่งหมายของการทดสอบ : เปรียบเทียบคุณภาพของการประเมิน

เวลาที่ใช้ในการทดสอบ 60 นาที ผู้วางแผน นางสาวทัตมณี ชูขวัญ

สาระการเรียนรู้	ผลการเรียนรู้ที่คาดหวัง	นำหนักรายข้อ	
		คิดเป็น%	จำนวนข้อ
1. ประโยค	1.1 ระบุประโยคที่เป็นจริงหรือประโยคที่เป็นเท็จได้	7.69	1
2. ศูนย์และจำนวนเต็มบวก	2.1 บอกสมบัติของศูนย์และหนึ่งได้	7.69	1
3. สมบัติการบวกและการคูณของจำนวนเต็มบวก	3.1 นำความรู้เกี่ยวกับสมบัติการสลับที่ สมบัติการเปลี่ยนกลุ่มและสมบัติการแจกแจงของการบวกและการคูณของจำนวนเต็มบวกไปแก้ปัญหาเกี่ยวกับการบวกและการคูณของจำนวนเต็มบวกได้	7.69	1
4. จำนวนเต็มลบและค่าสัมบูรณ์ของจำนวนเต็ม	4.1 บอกโครงสร้างของระบบจำนวนเต็มได้	7.69	1
	4.2 หาค่าสัมบูรณ์ของจำนวนเต็มได้	7.69	1
	4.3 หาจำนวนตรงข้ามของจำนวนเต็มได้	7.69	1
5. การบวกและการลบจำนวนเต็ม	5.1 เปลี่ยนรูปการบวกและการลบจำนวนเต็มที่ซับซ้อนให้อยู่ในรูปอย่างง่ายได้	7.69	1
6. การคูณและการหารจำนวนเต็ม	6.1 สรุปหลักการคูณและการหารจำนวนเต็มได้	7.69	1
	6.2 หาผลคูณและผลหารของจำนวนเต็มได้	7.69	1
	6.3 แก้ปัญหาการคูณและการหารจำนวนเต็มที่ซับซ้อนได้	7.69	1
7. สมบัติของหนึ่งและศูนย์	7.1 ระบุหลักการใช้สมบัติการคูณและการหารจำนวนใดด้วยหนึ่งได้	15.38	2
	7.2 นำสมบัติของหนึ่งและศูนย์ไปคำนวณหาผลบวกและผลคูณกับจำนวนเต็มได้	7.69	1
	รวม	100	13

ภาคผนวก ค

ค่าความยาก และค่าอำนาจจำแนก ของการวิเคราะห์ข้อสอบ ตามทฤษฎี CTT (ค่า pและ r) ของแบบทดสอบเรื่อง สมบัติของจำนวนนับ และระบบจำนวนเต็ม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 3 ค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบตามทฤษฎี CTT คือ ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบเรื่อง สมบัติของจำนวนนับ

ข้อที่	p	r	หมายเหตุ
1	0.82	0.30	ตัดทิ้ง
2	0.67	0.22	คัดเลือก
3	0.87	0.35	ตัดทิ้ง
4	0.73	0.33	คัดเลือก
5	0.60	0.54	คัดเลือก
6	0.87	0.30	ตัดทิ้ง
7	0.60	0.36	คัดเลือก
8	0.95	0.15	ตัดทิ้ง
9	0.82	0.18	ตัดทิ้ง
10	0.62	0.26	คัดเลือก
11	0.56	0.36	คัดเลือก
12	0.84	0.19	ตัดทิ้ง
13	0.78	0.34	คัดเลือก
14	0.73	0.55	คัดเลือก
15	0.95	0.10	ตัดทิ้ง
16	0.73	0.24	คัดเลือก
17	0.75	0.28	คัดเลือก
18	0.75	0.39	คัดเลือก
19	0.76	0.38	คัดเลือก
20	0.60	0.46	คัดเลือก
21	0.55	0.41	คัดเลือก
22	0.93	0.10	ตัดทิ้ง
23	0.80	0.13	ตัดทิ้ง
24	0.76	0.44	คัดเลือก
25	0.87	0.25	ตัดทิ้ง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางภาคผนวกที่ 4 ค่าพารามิเตอร์ของการวิเคราะห์ข้อสอบตามทฤษฎี CTT คือ ค่าความยาก (p) และค่าอำนาจจำแนก (r) ของแบบทดสอบเรื่อง สมบัติของจำนวนนับ

ข้อที่	p	r	หมายเหตุ
1	0.84	0.27	ตัดทิ้ง
2	0.49	0.55	คัดเลือก
3	0.85	0.33	ตัดทิ้ง
4	0.76	0.28	คัดเลือก
5	0.84	0.40	ตัดทิ้ง
6	0.62	0.67	คัดเลือก
7	0.73	0.40	คัดเลือก
8	0.71	0.28	คัดเลือก
9	0.55	0.35	คัดเลือก
10	0.29	0.43	คัดเลือก
11	0.69	0.41	คัดเลือก
12	0.55	0.61	คัดเลือก
13	0.64	0.41	คัดเลือก
14	0.78	0.15	คัดเลือก
15	0.75	0.40	คัดเลือก
16	0.42	0.36	คัดเลือก
17	0.49	0.55	คัดเลือก
18	0.49	0.68	คัดเลือก
19	0.69	0.41	คัดเลือก
20	0.71	0.02	ตัดทิ้ง
21	0.89	0.27	ตัดทิ้ง
22	0.55	0.68	คัดเลือก
23	0.76	0.60	คัดเลือก
24	0.45	0.56	คัดเลือก
25	0.85	0.20	ตัดทิ้ง

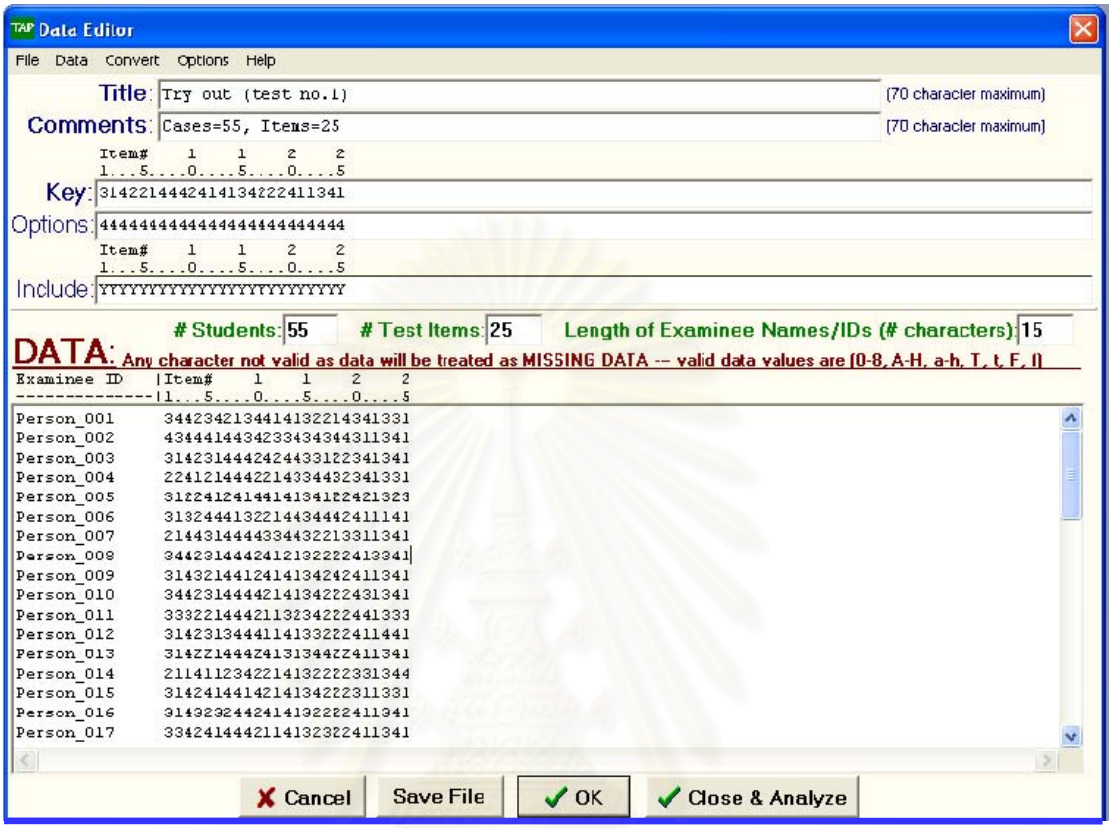
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

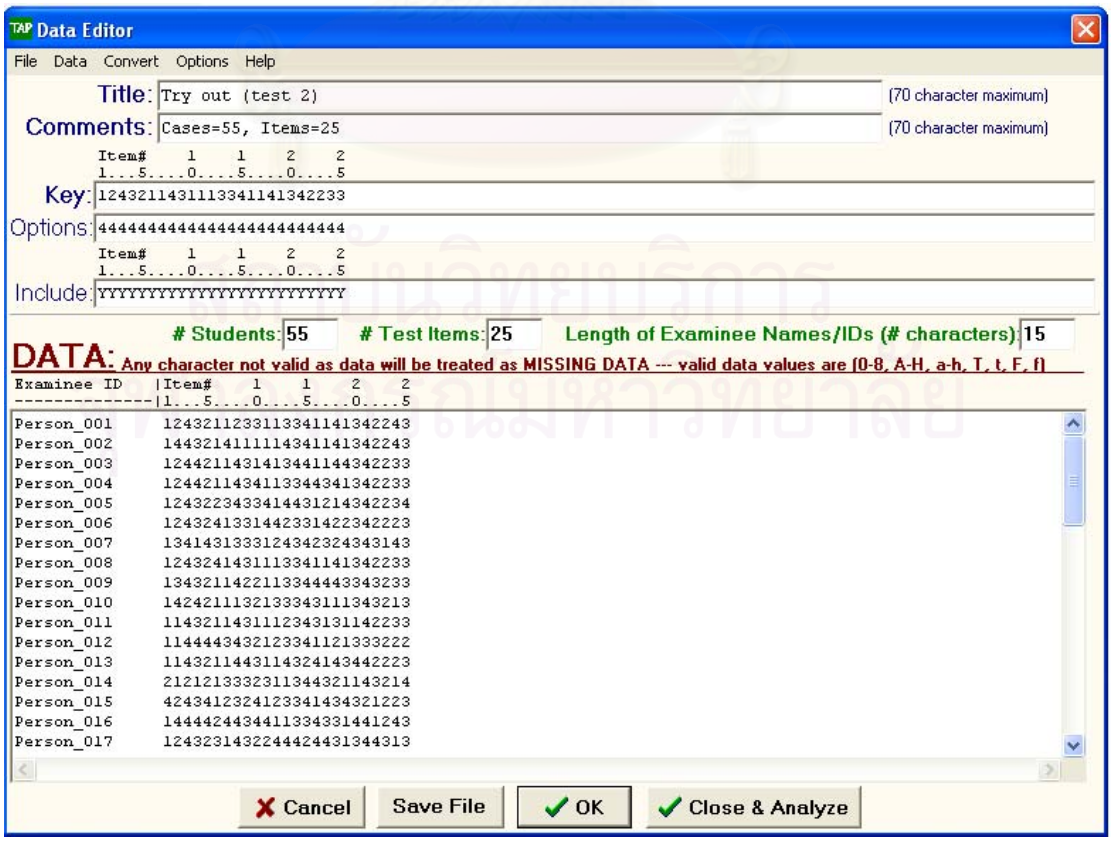
ภาคผนวก ง

การวิเคราะห์ข้อมูลโดยโปรแกรม TAP

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ข้อมูลโดยโปรแกรม TAP





ภาคผนวก จ

การวิเคราะห์ข้อมูลโดยโปรแกรม MULTILOG

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ข้อมูลโดยโปรแกรม MULTILOG

ตัวอย่าง File คำสั่ง

```
MULTILOG for Windows 7.00.2328.1
Created on: 12 April 2006, 23:40:51
>PROBLEM RANDOM,
         INDIVIDUAL,
         DATA = 'D:\MUL\NovakR1.DAT',
         NITEMS = 2,
         NGROUPS = 1,
         NEXAMINEES = 220,
         NCHARS = 3;
>TEST ALL,
      GRADED,
      NC = (8, 8);
>END ;
9
012345678
00
11
22
33
44
55
66
77
88
(3A1,2A1)
```

เมื่อ	NITEMS = 2	หมายถึง	จำนวนของข้อสอบ
	NGROUPS = 1,	หมายถึง	จำนวนกลุ่ม
	NEXAMINEES = 220,	หมายถึง	จำนวนผู้สอบ
	NCHARS = 3;	หมายถึง	จำนวนอักษรของรหัสผู้สอบ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ฉ

การวิเคราะห์ข้อมูลโดยโปรแกรม GENOVA

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ข้อมูลโดยโปรแกรม GENOVA

ตัวอย่าง File คำสั่ง

```
GSTUDY      P X I X R DESIGN -- RANDOM MODEL
OPTIONS     RECORDS 2
EFFECT      * P  220 0
EFFECT      + I  2 0
EFFECT      + R  2 0
FORMAT       (4F4.0)
PROCESS
   3   3   2   2 PERSON 1
   N   N   N   N
   3   3   3   2 PERSON 220
COMMENT     GSTUDY SET OF D STUDY CONTROL CARDS
DSTUDY      #1 -- P X I X R DESIGN -- RANDOM MODEL
DEFFECT     $ P
DEFFECT       I  2 / 2
DEFFECT       R  2
ENDDSTUDY
COMMENT     DSTUDY SET OF D STUDY CONTROL CARDS
DSTUDY      #2 -- P X I X R DESIGN -- RANDOM MODEL
DEFFECT     $ P
DEFFECT       I  4 8 12 16 20 24 28
DEFFECT       R  1
ENDDSTUDY
COMMENT     DSTUDY SET OF D STUDY CONTROL CARDS
DSTUDY      #3 -- P X I X R DESIGN -- RANDOM MODEL
DEFFECT     $ P
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
DEFFECT       I  4 8 12 16 20 24 28
DEFFECT       R  2
ENDDSTUDY
COMMENT     DSTUDY SET OF D STUDY CONTROL CARDS
DSTUDY      #4 -- P X I X R DESIGN -- RANDOM MODEL
DEFFECT     $ P
DEFFECT       I  4 8 12 16 20 24 28
DEFFECT       R  3
ENDDSTUDY
COMMENT     DSTUDY SET OF D STUDY CONTROL CARDS
DSTUDY      #5 -- P X I X R DESIGN -- RANDOM MODEL
DEFFECT     $ P
DEFFECT       I  4 8 12 16 20 24 28
DEFFECT       R  4
ENDDSTUDY
COMMENT     DSTUDY SET OF D STUDY CONTROL CARDS
DSTUDY      #6 -- P X I X R DESIGN -- RANDOM MODEL
DEFFECT     $ P
DEFFECT       I  4 8 12 16 20 24 28
DEFFECT       R  5
ENDDSTUDY
COMMENT     DSTUDY SET OF D STUDY CONTROL CARDS
DSTUDY      #7 -- P X I X R DESIGN -- RANDOM MODEL
DEFFECT     $ P
DEFFECT       I  4 8 12 16 20 24 28
DEFFECT       R  6
ENDDSTUDY
FINISH
```

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ช

คู่มือการใช้แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

คู่มือการใช้แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์

จุดมุ่งหมาย

แผนผังมโนทัศน์นี้สร้างขึ้นเพื่อใช้สำหรับประเมินความคิดรวบยอดวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 เพื่อทำให้ทราบถึงความคิดรวบยอด ที่ถูกต้องในเนื้อหาบทเรียนของนักเรียน จะได้เสริมสร้างและพัฒนากระบวนการคิดของนักเรียนได้อย่างเต็มศักยภาพ โดยแผนผังมโนทัศน์จะเป็นรูปแบบการประเมินทางเลือกใหม่สำหรับครูผู้สอนและนักเรียน

นิยามเชิงปฏิบัติการ

ความคิดรวบยอด หมายถึง ความรู้ ความเข้าใจที่เกิดจากกระบวนการคิดที่มีต่อเรื่องใดเรื่องหนึ่งซึ่งได้รับจากการเรียนรู้ หรือประสบการณ์ โดยสรุปออกมาเป็นถ้อยคำ กลุ่มคำหรือประโยคสั้น ๆ กะทัดรัด สามารถสื่อความหมายได้

แผนผังมโนทัศน์ หมายถึง แผนภาพหรือแผนภูมิที่แสดงกระบวนการคิดอย่างมีลำดับขั้นตอน โดยเสนอความสัมพันธ์ระหว่างมโนทัศน์หลายมโนทัศน์ ซึ่งจะเขียนมโนทัศน์หลักไว้ด้านบนสุดของแผนผังแล้ว เขียนมโนทัศน์รองหรือมโนทัศน์ย่อยลดหลั่นลงมา ลากลูกศรเชื่อมโยงระหว่างมโนทัศน์เพื่อแสดงทิศทางของความสัมพันธ์และเขียนคำหรือวลีกำกับไว้บนเส้นเพื่อแสดงความสัมพันธ์ระหว่างมโนทัศน์

ลักษณะของแบบประเมินความคิดรวบยอด

ลักษณะของแบบประเมินความคิดรวบยอดประกอบด้วย แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) โดยมีลักษณะดังนี้

แผนผังมโนทัศน์แบบสร้างผังจากคำ เป็นแบบประเมินกำหนดคำมโนทัศน์มาให้ แล้วให้นักเรียนสร้างแผนผังมโนทัศน์โดยใช้คำมโนทัศน์ที่กำหนดเหล่านั้น ซึ่งในการสร้างแผนผังมโนทัศน์จะให้อิสระแก่นักเรียน โดยจะไม่กำหนดรูปแบบของแผนผังนักเรียนสามารถสร้างได้ตามความ พอใจ นอกจากนี้นักเรียนสามารถเขียนคำมโนทัศน์ที่นอกเหนือคำมโนทัศน์ที่กำหนดให้เพิ่มลงในแผนผังได้ ในการสร้างนักเรียนจะต้องลากลูกศรเชื่อมโยงระหว่างมโนทัศน์ พร้อมทั้งเขียนคำเชื่อมเพื่ออธิบายความสัมพันธ์ระหว่างคำมโนทัศน์แต่ละคำ

โครงสร้างของแผนผังมโนทัศน์

แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์โดยใช้แผนผังมโนทัศน์ มีโครงสร้างเนื้อหาครอบคลุมสาระที่ 1 จำนวนและการดำเนินการ ช่วงชั้นที่ 3 ชั้นมัธยมศึกษาปีที่ 1 ดังนี้

สาระที่ 1 จำนวนและการดำเนินการ

มาตรฐาน	มาตรฐานการเรียนรู้ช่วงชั้น	ผลการเรียนรู้ที่คาดหวัง
มาตรฐานที่ ค1.1 เข้าใจถึงความหลากหลายของการแสดงจำนวนและการใช้จำนวนในชีวิตจริง	มีความคิดรวบยอดเกี่ยวกับจำนวนเต็มบวก เต็มลบ ศูนย์และจำนวนอตรรกยะ	● ระบุหรือยกตัวอย่างจำนวนเต็มบวก จำนวนเต็มลบ และศูนย์ได้ ● เปรียบเทียบจำนวนเต็มได้
มาตรฐานที่ ค1.2 เข้าใจถึงผลที่เกิดขึ้นจากการดำเนินการของจำนวนและความสัมพันธ์ระหว่างการดำเนินการต่างๆ และสามารถใช้การดำเนินการในการแก้ปัญหาได้	บวก ลบ คูณและหารจำนวนเต็มและนำไปใช้แก้ปัญหาได้	● บวก ลบ คูณและหารจำนวนเต็มได้
มาตรฐานที่ ค1.4 เข้าใจในระบบจำนวนและสามารถนำสมบัติเกี่ยวกับจำนวนไปใช้ได้	เข้าใจสมบัติต่างๆเกี่ยวกับระบบจำนวนเต็มและสามารถนำไปใช้แก้ปัญหาได้	● หา ห.ร.ม. และ ค.ร.น. ของจำนวนนับที่กำหนดให้ได้ ● ใช้ความรู้เกี่ยวกับ ห.ร.ม. และ ค.ร.น. แก้ปัญหาได้ ● นำความรู้และสมบัติเกี่ยวกับจำนวนเต็มไปใช้ได้

การตรวจให้คะแนน

1. การตรวจให้คะแนนแผนผังมโนทัศน์โดยใช้เกณฑ์การให้คะแนนของโนแวคและโกวิน

เกณฑ์การให้คะแนนของโนแวคและโกวิน จะพิจารณาแผนผังมโนทัศน์ในด้านคุณภาพ 4 ด้าน คือ พิจารณาจากประพจน์ (propositions) การจัดลำดับขั้นตอน (hierarchy) การเชื่อมระหว่างมโนทัศน์ (cross link) และตัวอย่าง (examples) ซึ่งแต่ละด้านจะแบ่งคะแนนออกเป็น 5 ระดับ จาก 0-4 ดังนั้นแผนผังมโนทัศน์ที่ตรวจให้ด้วยเกณฑ์การให้คะแนนของโนแวคและโกวินจะมีคะแนนเต็ม 16 คะแนน

ตารางภาคผนวกที่ 5 เกณฑ์การตรวจให้คะแนนของแผนผังมโนทัศน์ของโนแวคและโกวิน

ระดับคุณภาพ	ตัวบ่งชี้ ประพจน์ (propositions)
4	มโนทัศน์ที่เชื่อมกันตั้งแต่ 75% มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
3	มโนทัศน์ที่เชื่อมกันตั้งแต่ 50% แต่ไม่ถึง 75 % มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
2	มโนทัศน์ที่เชื่อมกันตั้งแต่ 25% แต่ไม่ถึง 50 % มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
1	มโนทัศน์ที่เชื่อมกันไม่ถึง 25 % มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์
0	ไม่พบมโนทัศน์ที่เชื่อมกันและคำเชื่อมมีความหมายและตรงกับความสัมพันธ์เลย
ระดับคุณภาพ	ตัวบ่งชี้ การจัดลำดับขั้นตอน (hierarchy)
4	แผนผังมโนทัศน์มีการจัดลำดับขั้นตอนตั้งแต่ 75% โดยแต่ละมโนทัศน์ที่อยู่รองลงมาเป็นมโนทัศน์ที่เฉพาะเจาะจง และมีความกว้างขวางน้อยกว่ามโนทัศน์ที่อยู่ลำดับแรก ๆ
3	แผนผังมโนทัศน์มีการจัด ลำดับขั้นตอนตั้งแต่ 50% แต่ไม่ถึง 75% โดยมีการจัดลำดับขั้นตอนเป็นส่วนมาก
2	แผนผังมโนทัศน์มีการจัด ลำดับขั้นตอนตั้งแต่ 25% แต่ไม่ถึง 50% โดยมีการจัดลำดับขั้นตอนเป็นบางส่วน
1	แผนผังมโนทัศน์มีการจัดลำดับขั้นตอนไม่ถึง 25% โดยมีการจัดลำดับขั้นตอนน้อยมาก
0	แผนผังมโนทัศน์ไม่มีการมีการจัดลำดับขั้นตอนเลย หรือการจัดลำดับขั้นตอนไม่ถูกต้องเลย
ระดับคุณภาพ	ตัวบ่งชี้ การเชื่อมระหว่างสายของมโนทัศน์ (cross link)
4	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์ตั้งแต่ 75% อย่างสมเหตุสมผลและถูกต้อง
3	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้องตั้งแต่ 50% แต่ไม่ถึง 75%
2	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้องตั้งแต่ 25% แต่ไม่ถึง 50%
1	มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้องไม่ถึง 25%
0	ไม่มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์อย่างสมเหตุสมผลและถูกต้อง

ตารางภาคผนวกที่ 5 เกณฑ์การตรวจให้คะแนนของแผนผังมโนทัศน์ของโนแวคและโกวิน (ต่อ)

ระดับคุณภาพ	ตัวบ่งชี้ ตัวอย่าง (example)
4	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องตั้งแต่ 75%
3	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องตั้งแต่ 50% แต่ไม่ถึง 75%
2	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องตั้งแต่ 25% แต่ไม่ถึง 50%
1	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้องไม่ถึง 25%
0	ไม่พบการยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้น

2. การตรวจให้คะแนนแผนผังมโนทัศน์โดยใช้เกณฑ์การให้คะแนนของบอลท์

เกณฑ์การให้คะแนนของบอลท์ จะพิจารณาให้คะแนนในด้านการจัดแผนผังมโนทัศน์ (organization) 6 คะแนน แบ่งคะแนนออกเป็น 7 ระดับ จาก 0-6 และความถูกต้อง (accuracy) 4 คะแนน แบ่งคะแนนออกเป็น 5 ระดับ จาก 0-4 ดังนั้นแผนผังมโนทัศน์ที่ตรวจให้ด้วยเกณฑ์การให้คะแนนของบอลท์จะมีคะแนนเต็ม 10 คะแนน

ตารางภาคผนวกที่ 6 เกณฑ์การตรวจให้คะแนนของแผนผังมโนทัศน์ของบอลท์

ระดับคุณภาพ	ตัวบ่งชี้ การจัดแผนผังมโนทัศน์
6	แผนผังมีความสมบูรณ์ โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กันทั้งหมด มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ 100%
5	แผนผังมีความสมบูรณ์เป็นส่วนใหญ่ โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ ได้ตั้งแต่ 80% แต่ไม่ถึง 100%
4	แผนผังมีความถูกต้องสมบูรณ์ปานกลาง มีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ได้ตั้งแต่ 60% แต่ไม่ถึง 80%

ตารางภาคผนวกที่ 6 เกณฑ์การตรวจให้คะแนนของแผนผังมโนทัศน์ของบอลท์ (ต่อ)

ระดับคุณภาพ	ตัวบ่งชี้ การจัดแผนผังมโนทัศน์
3	แผนผังมีความสมบูรณ์เป็นบางส่วน โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ได้ตั้งแต่ 40% แต่ไม่ถึง 80%
2	แผนผังมีความถูกต้องสมบูรณ์เล็กน้อย โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ ได้ตั้งแต่ 20% แต่ไม่ถึง 40%
1	แผนผังมีความถูกต้องสมบูรณ์น้อยมาก โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ได้ ใช้คำมโนทัศน์ที่กำหนดให้ได้น้อยกว่า 20%
0	แผนผังไม่มีความถูกต้องสมบูรณ์เลย การเชื่อมโยงมโนทัศน์ที่สัมพันธ์กันไม่ถูกต้อง ไม่ใช้ คำเชื่อมความสัมพันธ์ ไม่ยกตัวอย่างที่เป็นเหตุการณ์หรือวัตถุที่เป็นตัวแทนของมโนทัศน์ ไม่ได้ใช้คำมโนทัศน์ที่กำหนดให้
ระดับคุณภาพ	ตัวบ่งชี้ ความถูกต้อง
4	ไม่ผิดพลาดเลย
3	ผิดในรายละเอียดเล็กน้อย
2	คลาดเคลื่อนเล็กน้อยและเป็นการคลาดเคลื่อนในมโนทัศน์รอง
1	คลาดเคลื่อนในมโนทัศน์หลักเป็นส่วนใหญ่
0	ไม่ถูกต้องเลย หรือคลาดเคลื่อนในมโนทัศน์หลักหลายมโนทัศน์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์

คำชี้แจง

แบบประเมินความคิดรวบยอดวิชาคณิตศาสตร์ ประกอบด้วยแผนผังมโนทัศน์แบบสร้างผังจากคำ (construct a map) โดยมีเนื้อหาครอบคลุม สาระที่ 1 จำนวนและการดำเนินการ ชั้นมัธยมศึกษาปีที่ 1 ซึ่งแบบประเมินมีทั้งหมด 2 ชุด

ชุดที่ 1 เรื่อง สมบัติของจำนวนนับ

ชุดที่ 2 เรื่อง ระบบจำนวนเต็ม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แบบประเมินชุดที่ 1 เรื่อง สมบัติของจำนวนนับ

แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map)

คำชี้แจง

1) แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) เรื่อง สมบัติของจำนวนนับ
2) ให้นักเรียนพิจารณาคำมโนทัศน์ในกรอบสี่เหลี่ยมด้านล่าง แล้วนำคำมโนทัศน์มาสร้างเป็นแผนผัง
3) วิธีการสร้างแผนผังมโนทัศน์

ให้นักเรียนเลือกคำมโนทัศน์หลักหรือมโนทัศน์สำคัญเขียนวางไว้ด้านบนสุดของหน้ากระดาษ จากนั้นเขียนมโนทัศน์ย่อยหรือมโนทัศน์รองลงมาเรื่อย ๆ จนครบทุกคำโดยสามารถวาดเป็นรูปสี่เหลี่ยม วงกลม หรือวงรี ล้อมรอบคำได้ตามความต้องการ จากนั้นลากเส้นเชื่อมโยงระหว่างมโนทัศน์ พร้อมทั้งเขียนคำเชื่อมหรือวลีกำกับไว้บนเส้นเพื่อแสดงถึงความสัมพันธ์ระหว่างมโนทัศน์ โดยในการสร้างแผนผังนักเรียนสามารถเพิ่มคำมโนทัศน์ที่เกี่ยวข้องอื่น ๆ เข้าไปได้ตามความต้องการ

4) เมื่อนักเรียนสร้างแผนผังเสร็จแล้วให้ตรวจสอบดังนี้
   4.1 นักเรียนใช้คำมโนทัศน์ที่กำหนดให้ครบทุกคำ
   4.2 นักเรียนลากเส้นเชื่อมโยงระหว่างมโนทัศน์พร้อมทั้งเขียนคำเชื่อมเพื่อแสดงความสัมพันธ์ระหว่างมโนทัศน์ครบทุกเส้น

<u>คำมโนทัศน์ เรื่อง สมบัติของจำนวนนับ</u>

จำนวนนับ ตัวประกอบ ตัวประกอบร่วม ตัวหารร่วม ตัวหารร่วมมาก
ขั้นตอนวิธีการหารแบบยูคลิด วิธีการหารด้วยจำนวนเฉพาะ ตัวคูณร่วมน้อย
พหุคูณ การแยกตัวประกอบเฉพาะ จำนวนเฉพาะ จำนวนเต็ม
จำนวนคู่ จำนวนคี่

สมบัติของจำนวนนับ	

ชื่อ_____ชั้น_____เลขที่_____

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แบบประเมินชุดที่ 2 เรื่อง ระบบจำนวนเต็ม

แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map)

คำชี้แจง

1) แผนผังมโนทัศน์แบบสร้างผังจากคำ (construct-a-map) เรื่อง ระบบจำนวนเต็ม
2) ให้นักเรียนพิจารณาคำมโนทัศน์ในกรอบสี่เหลี่ยมด้านล่าง แล้วนำคำมโนทัศน์มาสร้างเป็นแผนผัง
3) วิธีการสร้างแผนผังมโนทัศน์

ให้นักเรียนเลือกคำมโนทัศน์หลักหรือมโนทัศน์สำคัญเขียนวางไว้ด้านบนสุดของหน้ากระดาษ จากนั้นเขียนมโนทัศน์ย่อยหรือมโนทัศน์รองลงมาเรื่อย ๆ จนครบทุกคำโดยสามารถวาดเป็นรูปสี่เหลี่ยม วงกลม หรือวงรี ล้อมรอบคำได้ตามความต้องการ จากนั้นลากเส้นเชื่อมโยงระหว่างมโนทัศน์ พร้อมทั้งเขียนคำเชื่อมหรือวลีกำกับไว้บนเส้นเพื่อแสดงถึงความสัมพันธ์ระหว่างมโนทัศน์ โดยในการสร้างแผนผังนักเรียนสามารถเพิ่มคำมโนทัศน์ที่เกี่ยวข้องอื่น ๆ เข้าไปได้ตามความต้องการ

4) เมื่อนักเรียนสร้างแผนผังเสร็จแล้วให้ตรวจสอบดังนี้
   4.1 นักเรียนใช้คำมโนทัศน์ที่กำหนดให้ครบทุกคำ
   4.2 นักเรียนลากเส้นเชื่อมโยงระหว่างมโนทัศน์พร้อมทั้งเขียนคำเชื่อมเพื่อแสดงความสัมพันธ์ระหว่างมโนทัศน์ครบทุกเส้น

<u>คำมโนทัศน์ เรื่อง ระบบจำนวนเต็ม</u>

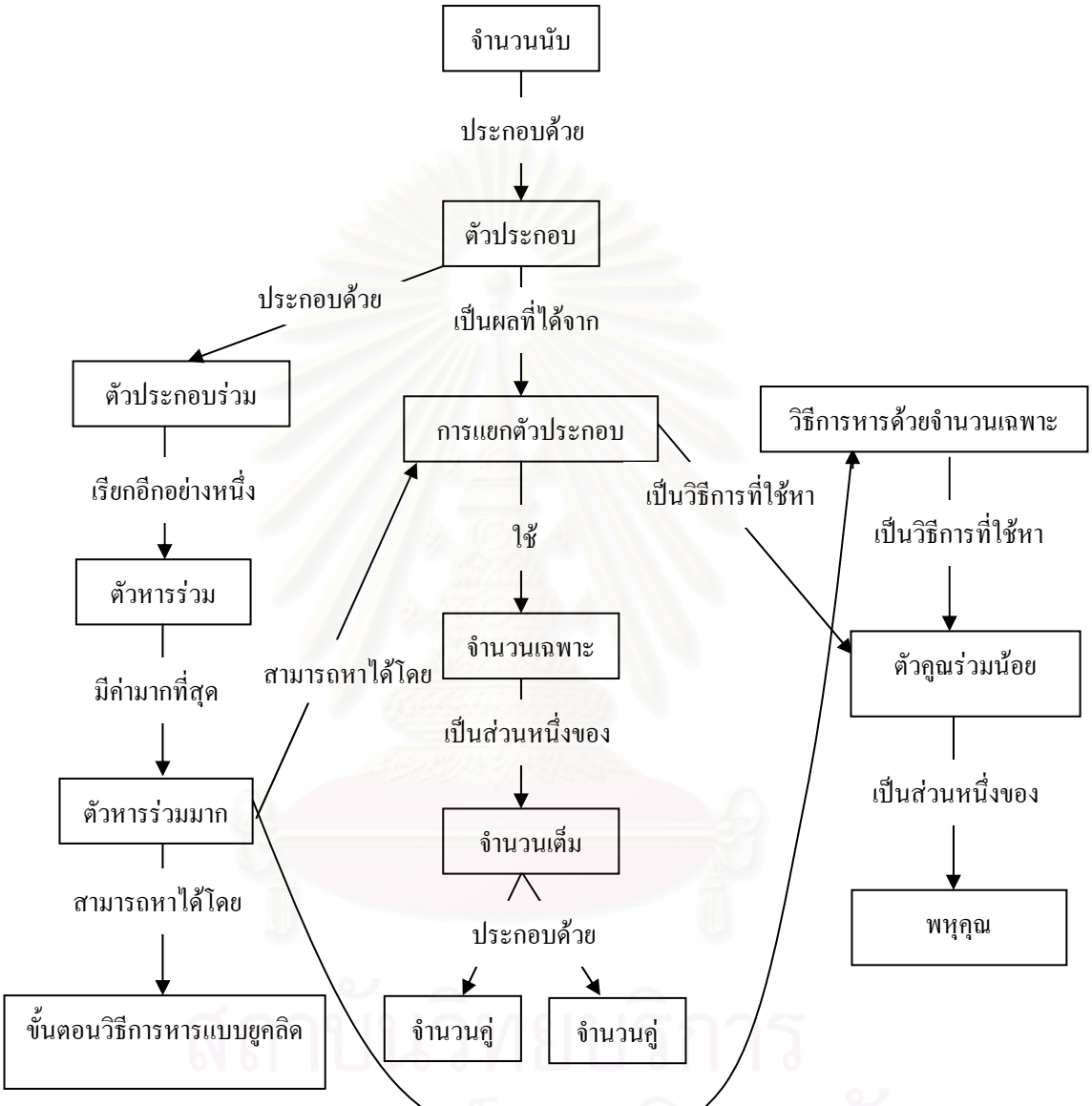
จำนวนเต็ม	จำนวนเต็มลบ	ศูนย์	จำนวนเต็มบวก
จำนวนนับ	ค่าสัมบูรณ์	จำนวนตรงข้าม	สมบัติการสลับที่
สมบัติการแจกแจง	สมบัติการเปลี่ยนกลุ่ม	การบวก	การคูณ
เอกลักษณ์			

ระบบจำนวนเต็ม	

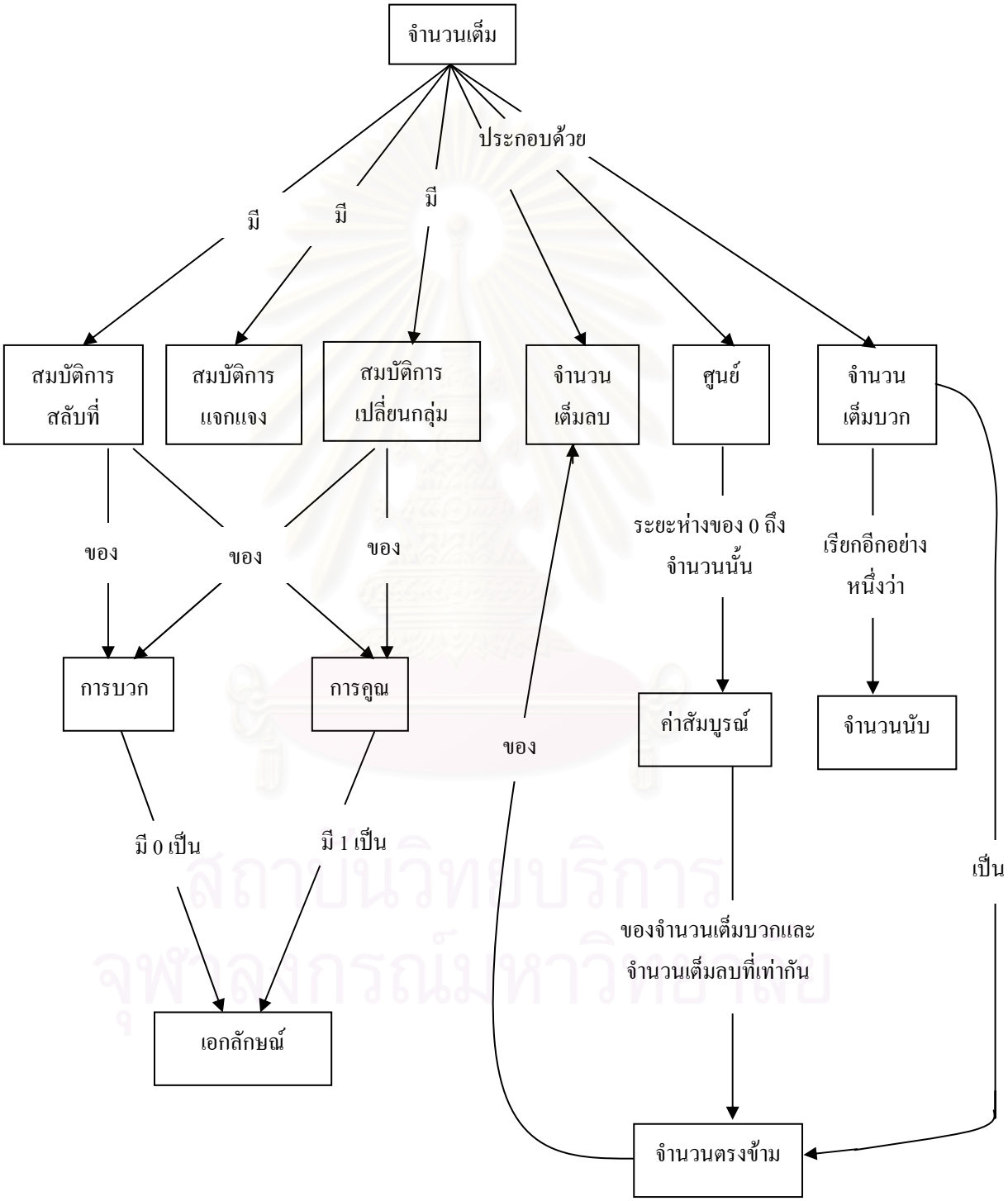
ชื่อ_____ชั้น_____เลขที่_____

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

แผนผังเกณฑ์ ชุดที่ 1 เรื่องสมบัติของจำนวนนับ



แผนผังเกณฑ์ ชุดที่ 2 เรื่องระบบจำนวนเต็ม



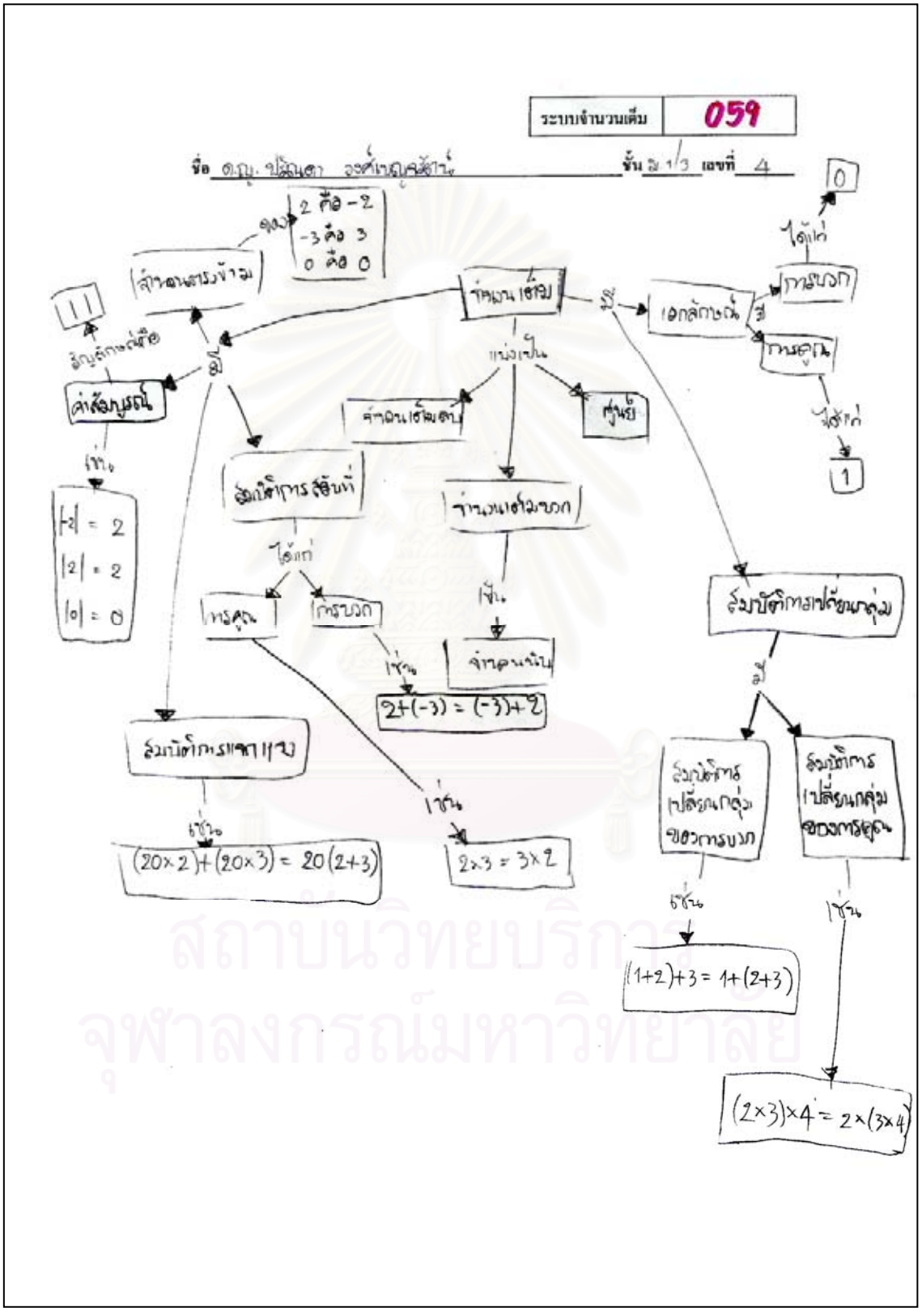
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ซ

ตัวอย่างผลงานของนักเรียนและการตรวจให้คะแนน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตัวอย่างผลงานของนักเรียน



ตัวอย่างการตรวจให้คะแนน

เรื่อง ระบบจำนวนเต็ม เลขที่ 59 (ภาคผนวก ซ หน้า 155)

ชื่อ เด็กหญิง ปริณดา วงศ์เบญจรัตน์ ชั้น ม.1/3 เลขที่ 4

การตรวจให้คะแนนด้วยวิธีของโนแวคและโกวิน

ตัวบ่งชี้	การตอบของนักเรียน	คะแนน
ประพจน์ (propositions)	มโนทัศน์ที่เชื่อมกัน 100% มีคำเชื่อมที่ความหมายและตรงกับความสัมพันธ์	4
การจัดลำดับขั้นตอน (hierarchy)	แผนผังมโนทัศน์มีการจัดลำดับขั้นตอนตั้งแต่ 80% โดยแต่ละมโนทัศน์ที่อยู่รองลงมาเป็นมโนทัศน์ที่เฉพาะเจาะจง และมีความกว้างขวางน้อยกว่ามโนทัศน์ที่อยู่ลำดับแรก ๆ	4
การเชื่อมระหว่างสายของมโนทัศน์ (cross link)	ไม่มีการเชื่อมโยงกันระหว่างสายของมโนทัศน์	0
ตัวอย่าง (example)	ยกตัวอย่างของเหตุการณ์หรือวัตถุที่เป็นตัวแทนของ มโนทัศน์ที่สร้างขึ้นได้ถูกต้อง 100%	4
รวม		**12**

การตรวจให้คะแนนด้วยวิธีของบอลท์

ตัวบ่งชี้	การตอบของนักเรียน	คะแนน
การจัดแผนผังมโนทัศน์ (organization)	แผนผังมีความสมบูรณ์เป็นส่วนใหญ่ โดยมีการเชื่อมโยงมโนทัศน์ที่สัมพันธ์กัน มีการใช้คำเชื่อมความสัมพันธ์ ใช้คำมโนทัศน์ที่กำหนดให้ ทั้งหมด แต่ยังขาดการยกตัวอย่าง	5
ความถูกต้อง (accuracy)	ผิดในรายละเอียดเล็กน้อย	3
รวม		**8**

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวทัตมณี ชูขวัญ เกิดเมื่อวันที่ 11 พฤษภาคม พ.ศ. 2523 สำเร็จการศึกษาหลักสูตรปริญญาครุศาตรบัณฑิต สาขามัธยมศึกษา-คณิตศาสตร์ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2545 และเข้าศึกษาต่อในหลักสูตรครุศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2547

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย