

การศึกษาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสม
เมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน



นางสาวปวีณา ป้ออาทิตย์


วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรมหาบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2545

ISBN 974-17-0876-9

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A STUDY OF THE OPTIMAL NUMBER OF RATERS
AND NUMBER OF WRITING TASKS
USING DIFFERENT SCORING RUBRICS



Miss Paweena Peeartit

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Education in Educational Measurement and Evaluation

Department of Educational Research

Faculty of Education

Chulalongkorn University

Academic Year 2002

ISBN 974-17-0876-9

ปีถัดมา ปีอาทิตย์: การศึกษาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้
เกณฑ์การให้คะแนนที่ต่างกัน (A STUDY OF THE OPTIMAL NUMBER OF RATERS
AND NUMBER OF WRITING TASKS USING DIFFERENT SCORING RUBRICS)
อ.ที่ปรึกษา:รศ. พวงแก้ว ปุณยกนก, 94 หน้า. ISBN 974-17-0876-9

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) เปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง
ของคะแนนการประเมินงานเขียนจากผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic
scoring rubric) และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring
rubric) และ 2) เพื่อหาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพ
รวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)
กลุ่มตัวอย่างประกอบด้วย ผลงานการเขียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 จำนวน 117
ผลงาน และผู้ประเมินจำนวน 8 คน ที่ใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยก
องค์ประกอบ การวิจัยนี้คำนวณค่าความเที่ยงจากค่าสัมประสิทธิ์การสรุปอ้างอิงด้วยโปรแกรมสำเร็จรูป
GENOVA

ผลการวิจัยพบว่า

1. ภายใต้ชิ้นงานจำนวน 117 ชิ้น ผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม
ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.71484 และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ
ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.81410
2. ถ้าต้องการใช้ผู้ประเมิน 1 คน และกำหนดให้ค่าสัมประสิทธิ์การสรุปอ้างอิงของเกณฑ์การให้
คะแนนแบบภาพรวมเป็น 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 ใช้จำนวนชิ้นงานอย่างน้อย 12 ชิ้น ทั้งนี้ขึ้นอยู่กับผู้ปฏิบัติ
ว่าจะเลือกใช้เกณฑ์การให้คะแนนให้เหมาะสมกับการเรียนการสอน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา ศึกษาศาสตร์
สาขาวิชา การวัดและประเมินผลการศึกษา
ปีการศึกษา 2545

ลายมือชื่อนิสิต.....
ลายมือชื่ออาจารย์ที่ปรึกษา
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม —

438 37145 27: MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION

KEY WORD: HOLISTIC/ANALYTIC/RATER/SCORING RUBRIC/GENERALIZABILITY COEFFICIENT

PAWEENA PEEARTIT: A STUDY OF THE OPTIMAL NUMBER OF RATERS AND
NUMBER OF WRITING TASKS USING DIFFERENT SCORING RUBRICS .

THESIS ADVISOR: ASSOC. PROF. PUANGKAEW POONYAKANOK, 94 pp.

ISBN 974-17-0876-9

The purposes of this research were 1) to compare the reliability from generalizability coefficient of writing task assessment based on 4 raters who used holistic scoring rubric and 4 raters who used analytic scoring rubric 2) to study the optimal number of raters and the number of writing tasks when using holistic scoring rubric or analytic scoring rubric. The sample consisted of 117 students' writing in Thai subject, Phathomsuksa V and 8 raters who scored tasks by using different scoring methods; holistic scoring rubric and analytic scoring rubric. The reliability of scoring was analysed by using GENOVA program.

The research findings were as follows:

1. For 117 writing tasks and 4 raters, the Generalizability Coefficients of holistic scoring rubric was 0.71484 and the Generalizability Coefficients of analytic scoring rubric was 0.81410 .
2. If using 1 rater and a Generalizability Coefficient of holistic scoring rubric was set at 0.2, 0.4, 0.6 and 0.8 the number of writings should be more than 4, 8, 16 and 28 pieces respectively. If using 1 rater and a Generalizability Coefficient of analytic scoring rubric was set at 0.2 and 0.4 the number of writings should be less than 4 pieces. If a Generalizability Coefficient was set at 0.6 the number of writings should be more than 8 pieces. If a Generalizability Coefficient was set at 0.8 the number of writings should be more than 12 pieces. Selection appropriate scoring rubric for instruction depends on user.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department Educational Research Student's signature.....

Field of study Educational Measurement and Evaluation Advisor's signature

Academic year 2002 Co-advisor's signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดีด้วยความเมตตากรุณาอย่างยิ่งจาก รองศาสตราจารย์พวงแก้ว ปุณยกันก อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่คอยให้คำแนะนำ คำปรึกษา ความคิดเห็น ตรวจสอบแก้ไขข้อบกพร่อง และคอยให้กำลังใจผู้วิจัยตลอดเวลาตั้งแต่เริ่ม ค้นคว้าเอกสาร ทำให้วิทยานิพนธ์ฉบับนี้มีความสมบูรณ์มากที่สุด ผู้วิจัยรู้สึกซาบซึ้งและขอกราบ ขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอกราบขอบพระคุณคณาจารย์ในภาควิชาวิจัยการศึกษาทุกท่านที่ได้ประสิทธิ์ประสาท ความรู้วิทยาการต่างๆให้แก่ผู้วิจัย

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ ผู้ทรงคุณวุฒิทุกท่านที่ท่านได้สละเวลาในการตรวจสอบเครื่องมือและขอขอบพระคุณผู้บริหารโรงเรียนและอาจารย์ผู้ปฏิบัติงานทุกท่านที่ให้ความร่วมมือในการเก็บรวบรวมข้อมูลเป็นอย่างดี รวมทั้งขอขอบคุณบัณฑิตวิทยาลัยที่ให้ ทุนอุดหนุนในการทำวิจัยในครั้งนี้

ขอขอบพระคุณอาจารย์กนกวรรณ เขี่ยมชัย ที่อนุเคราะห์โปรแกรมและคู่มือการใช้ โปรแกรม GENOVA และให้คำปรึกษา แนะนำ ช่วยเหลือในการใช้โปรแกรม

ขอขอบคุณทุกแรงกำลังใจจากเพื่อนๆ พี่ๆ ทุกท่านที่ให้กำลังใจเสมอมาและให้ความช่วยเหลือในด้านต่างๆในการทำวิทยานิพนธ์

ขอขอบคุณผู้ที่มีส่วนช่วยเหลือทุกท่าน ที่ไม่สามารถเอ่ยนามในที่นี้ได้ทั้งหมดที่มีส่วนช่วย ในการทำวิทยานิพนธ์จนสำเร็จทุกขั้นตอน

ท้ายที่สุดนี้ ผู้วิจัยขอขอบคุณค่าของงานวิจัยนี้ และความดีทั้งหมดแก่คุณพ่อสมนึก คุณแม่สมถวิล ป้าอาทิตย์ ผู้ที่คอยให้ความรัก กำลังกาย กำลังใจ กำลังทรัพย์ ด้วยความรักและห่วงใยมากที่สุด และส่งเสริมผู้วิจัยให้ได้รับการศึกษามาจนถึงระดับนี้อย่างที่หาคุณความดีอื่นใด เปรียบมิได้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ปวีณา ป้าอาทิตย์

สารบัญ

หน้า

| | |
|-------------------------|---|
| บทคัดย่อภาษาไทย..... | ง |
| บทคัดย่อภาษาอังกฤษ..... | จ |
| กิตติกรรมประกาศ..... | ฉ |
| สารบัญ..... | ช |
| สารบัญตาราง..... | ฅ |
| สารบัญภาพ..... | ฉ |

บทที่

| | |
|---|----|
| 1 บทนำ..... | 1 |
| ความเป็นมาและความสำคัญของปัญหา..... | 1 |
| วัตถุประสงค์การวิจัย..... | 6 |
| ขอบเขตการวิจัย..... | 6 |
| คำจำกัดความที่ใช้ในการวิจัย..... | 7 |
| ข้อตกลงเบื้องต้น..... | 8 |
| ประโยชน์ที่คาดว่าจะได้รับ..... | 8 |
| 2 เอกสารและงานวิจัยที่เกี่ยวข้อง..... | 9 |
| ตอนที่ 1 การประเมินแบบดั้งเดิมและการประเมินตามสภาพจริง..... | 9 |
| ตอนที่ 2 เกณฑ์การให้คะแนน..... | 14 |
| ตอนที่ 3 ทฤษฎีการสรุปอ้างอิง..... | 21 |
| ตอนที่ 4 งานวิจัยอื่นๆที่เกี่ยวกับเกณฑ์การให้คะแนนและทฤษฎีการสรุปอ้างอิง..... | 30 |
| 3 วิธีดำเนินการวิจัย..... | 34 |
| ประชากรและกลุ่มตัวอย่าง..... | 34 |
| การสร้างและพัฒนาเครื่องมือ..... | 36 |
| การเก็บรวบรวมข้อมูล..... | 43 |
| การวิเคราะห์ข้อมูล..... | 44 |

สารบัญ (ต่อ)

| | หน้า |
|---|------|
| 4 ผลการวิเคราะห์ข้อมูล..... | 46 |
| สถิติพื้นฐาน..... | 47 |
| การหาค่าความเที่ยงของคะแนนการประเมินงานเขียน..... | 47 |
| 4.1 ผลการประมาณค่าความแปรปรวน..... | 47 |
| 4.2 การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง..... | 50 |
| 4.3 การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง..... | 50 |
| 5 สรุปผล อภิปรายและข้อเสนอแนะ..... | 61 |
| สรุปผลการวิจัย..... | 62 |
| อภิปรายผลการวิจัย..... | 65 |
| ข้อเสนอแนะในการนำผลการวิจัยไปใช้..... | 68 |
| ข้อเสนอแนะในการวิจัยครั้งต่อไป..... | 68 |
| รายการอ้างอิง..... | 69 |
| ภาคผนวก..... | 73 |
| ภาคผนวก ก เครื่องมือที่ใช้ในการวิจัย..... | 74 |
| ภาคผนวก ข ตัวอย่างผลงานนักเรียน..... | 81 |
| ภาคผนวก ค ผลการตรวจผลงานการเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ค่าสถิติพื้นฐาน..... | 85 |
| ภาคผนวก ง รายนามผู้ทรงคุณวุฒิที่ตรวจสอบเกณฑ์การให้คะแนน..... | 90 |
| ภาคผนวก จ รายนามอาจารย์ที่ตรวจให้คะแนนผลงาน..... | 92 |
| ประวัติผู้เขียนวิทยานิพนธ์..... | 94 |

สารบัญตาราง

| ตาราง | หน้า |
|---|------|
| 1 การประมาณค่าความแปรปรวนจากแหล่งต่างๆของ PXIXR Design..... | 45 |
| 2 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมิน จำแนกตามผู้ประเมิน และตามรูปแบบเกณฑ์การให้คะแนน..... | 47 |
| 3 ผลการประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อ ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ($n_p=39, n_i=3, n_r=4$)..... | 48 |
| 4 ผลการประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อ ค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ($n_p=39, n_i=3, n_r=4$)..... | 49 |
| 5 ผลการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมและแบบแยกองค์ประกอบ..... | 50 |
| 6 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินหนึ่งคน..... | 51 |
| 7 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินสองคน..... | 52 |
| 8 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินสามคน..... | 53 |
| 9 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินสี่คน..... | 54 |
| 10 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินห้าคน..... | 55 |
| 11 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินหกคน..... | 56 |

สารบัญตาราง (ต่อ)

| ตาราง | หน้า |
|---|------|
| 12 ผลการตรวจผลงานการเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ค่าสถิติพื้นฐาน..... | 86 |
| 13 ผลการตรวจผลงานการเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ค่าสถิติพื้นฐาน..... | 88 |



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

| ภาพที่ | หน้า |
|--------|--|
| 1 | แผนภูมิแสดงองค์ประกอบความแปรปรวนของรูปแบบ PXIXR เมื่อตัวประกอบ ที่ศึกษาทั้งหมดเป็นตัวประกอบสุ่ม.....26 |
| 2 | แสดงขั้นตอนการตรวจผลงานของผู้ประเมิน.....35 |
| 3 | แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมิน งานเขียนที่ได้จากการใช้เกณฑ์การให้คะแนนแบบภาพรวม ที่จำนวนงานเขียน 4-28 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามแบบแผนการวัด P X I X R.....57 |
| 4 | แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมิน งานเขียนที่ได้จากการใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบที่จำนวน งานเขียน 4-28 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามแบบแผนการวัด P X I X R.....58 |

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในสภาพของโลกปัจจุบันที่มีการเปลี่ยนแปลงอย่างรวดเร็วทำให้สภาพการเรียนการสอนปัจจุบันที่เน้นความรู้ ความจำ คงไม่เพียงพอที่จะพัฒนาคนให้สามารถดำรงชีวิตอยู่ในสังคมได้อย่างปกติสุข จากพระราชบัญญัติการศึกษาแห่งชาติ 2542 ในหมวด 4 แนวการจัดการสอน มาตรา 22 กล่าวว่า “การจัดการการศึกษาต้องยึดหลักว่าผู้เรียนทุกคนมีความสามารถเรียนรู้และพัฒนาตนเองได้และถือว่าผู้เรียนมีความสำคัญที่สุด กระบวนการจัดการการศึกษาต้องส่งเสริมให้ผู้เรียนสามารถพัฒนาตามธรรมชาติและเต็มทางศักยภาพ” และในมาตรา 26 กล่าวว่า “ให้สถานศึกษาจัดการประเมินผู้เรียนโดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติสังเกตพฤติกรรมการเรียน การร่วมกิจกรรมและการทดสอบควบคู่ไปในกระบวนการเรียนการสอนตามความเหมาะสมของแต่ละระดับและรูปแบบการศึกษา” ดังนั้นจึงทำให้เกิดการเปลี่ยนแปลงในเรื่องของกระบวนการเรียนรู้ วิธีการเรียนรู้ การจัดการเรียนโดยเน้นผู้เรียนเป็นสำคัญ และมีการผสมผสานการเรียนการสอนและการวัดผลประเมินผลไว้ด้วยกัน

การประเมินผลทางการศึกษาที่ปฏิบัติกันมาจนเป็นประเพณีนิยม มักอาศัยผลการสอบเขียนตอบหรือเลือกคำตอบเป็นหลัก มีการสอบภาคปฏิบัติบ้างขึ้นอยู่กับประสบการณ์ความพร้อมของครู ในทฤษฎีการทดสอบถือว่าข้อสอบแต่ละฉบับเป็นเพียงตัวอย่างเนื้อหาที่นำมาประกอบกับพฤติกรรมที่ต้องการตรวจสอบเท่านั้น ข้อสอบเป็นเพียงตัวอย่างคำถามที่ใช้กระตุ้นให้ผู้สอบแสดงพฤติกรรมที่ต้องการตรวจสอบและที่สำคัญที่สุด คือ การทดสอบทุกครั้งจะเกิดความคลาดเคลื่อนในการวัด ซึ่งจะต้องนำมาพิจารณาในการแปลความหมายคะแนนที่แต่ละคนทำได้ด้วย ปัจจุบันนี้ยอมรับกันว่าการสร้างข้อสอบและการทดสอบของครูมีจุดอ่อนที่ก่อให้เกิดความคลาดเคลื่อนของผลการวัดอย่างมาก คะแนนหรือผลการสอบเป็นข้อมูลที่สะท้อนความจริงบางส่วนของผู้เรียนเท่านั้น นอกจากนี้ยังมีพฤติกรรมอีกจำนวนมากที่ไม่สามารถตรวจสอบด้วยการใช้ข้อสอบให้ผู้เรียนเขียนตอบหรือเลือกตอบ ซึ่งกระบวนการเหล่านี้มีข้อจำกัดในการใช้เพื่อปรับปรุงแก้ไขจุดอ่อนในด้านต่างๆของผู้เรียนเป็นรายบุคคล บ่อยครั้งพบว่าผู้เรียนไม่สามารถประมาณความสามารถของตนและไม่ทราบว่ามีมาตรฐานของการปฏิบัติที่ควรจะเป็นตามความคาดหวังของหลักสูตรหรือผู้สอนเป็นอย่างไร ครูเองขาดทิศทางของการจัดกิจกรรมการสอนที่ให้โอกาสนักเรียนได้พัฒนาตนเองอย่างเต็มที่ (ภาวิณี ศรีสุขวัฒนานนท์, 2540)

ในปี ค.ศ.1986 สำนักพิมพ์ Psychological Corporation ในสหรัฐอเมริกาซึ่งเป็นสำนักพิมพ์แบบทดสอบมาตรฐานและตำราได้พิมพ์คำขวัญในหนังสือกำหนดการประชุมประจำปีของสมาคมวิจัยทางการศึกษาไว้ว่า “ในเนื้อผ้าแห่งการศึกษา การสอนและการประเมินจะถูกทักถอดประสานกันทำให้เนื้อผ้าแข็งแรงและทั้งสองส่วนนี้จะช่วยให้เราเข้าถึงสมรรถภาพสูงสุดของเรา (In the fabric of education, instruction and assessment are interwoven. Each strengthens the others. Together they enable us to realize our full potential)” จากคำกล่าวข้างต้นแสดงให้เห็นว่า การเรียนการสอนและการประเมินไม่สามารถแยกส่วนออกจากกันได้ เมื่อครูกำหนดกิจกรรมการเรียนรู้ของนักเรียน โดยที่นักเรียนเป็นผู้ทำกิจกรรมนี้ ครูจะต้องพร้อมที่จะประเมินผลการปฏิบัติกิจกรรมของนักเรียนด้วย แต่พบว่าปัญหาสำคัญประการหนึ่งของการวัดและประเมินผลการเรียนในปัจจุบันคือ ครูผู้สอนมักแยกการเรียนการสอนและการวัดประเมินผลออกจากกัน ซึ่งแท้จริงแล้วการเรียนการสอน และการวัดประเมินผลควรดำเนินไปด้วยกัน และเป็นไปอย่างต่อเนื่องมีความสัมพันธ์ซึ่งกันและกัน และเกิดขึ้นในเวลาเดียวกัน และปัญหาที่สำคัญอีกประการหนึ่งคือ ครูผู้สอนใช้แบบทดสอบในการวัด และประเมินผลเป็นส่วนใหญ่หรือกล่าวว่าเป็นเครื่องมือเดียวที่ผู้สอนใช้ ซึ่งแบบทดสอบมีข้อจำกัด ข้อบกพร่องหลายประการ ในการใช้ประเมินหรือตัดสินเกี่ยวกับตัวผู้เรียนเนื่องจากแบบทดสอบปรนัยวัดได้เพียงความรู้ ความจำ ซึ่งไม่ครอบคลุมพฤติกรรมทุกด้านที่ต้องการให้ผู้เรียนเกิด นอกจากนี้ยังไม่สามารถวัดกระบวนการความคิดที่ซับซ้อนหรือความคิดระดับสูงของผู้เรียนได้ วัดประสิทธิภาพการแก้ปัญหาไม่ได้ วัดกระบวนการเรียนของผู้เรียนได้ค่อนข้างน้อย วัดทักษะบางอย่างไม่ได้ เช่น การพูด การเขียน การปฏิบัติ การสร้างสรรค์ การเล่นดนตรี เป็นต้น

ด้วยเหตุนี้เองจึงต้องมีการปฏิรูปและหาแนวทางเลือกใหม่ที่เหมาะสมในการประเมินผล โดยเปิดโอกาสให้ผู้เรียนได้ลงมือปฏิบัติจริง เพราะการปฏิบัติผู้เรียนจำเป็นต้องใช้ความรู้ และทักษะที่มีอยู่เดิมมาคิดแก้ปัญหาใหม่ ซึ่งเป็นวิธีการเรียนรู้ที่สอดคล้องกับสภาพจริงของการดำรงชีวิตในสังคม ดังนั้นรูปแบบการประเมินรูปแบบใหม่ ควรมีรูปแบบดังนี้ (สมชาย มิ่งมิตร, 2539)

1. เน้นให้นักเรียนตอบด้วยการแสดง สร้างสรรค์ ผลิต หรือทำงาน
2. ดึงเอาความคิดขั้นสูง ความคิดซับซ้อน และทักษะการแก้ปัญหาออกมาได้
3. ใช้งานที่เป็นผลจากการเรียนการสอนที่สอดคล้องกับความเป็นจริงในชีวิตประจำวัน
4. กระตุ้นให้เกิดการประยุกต์สู่โลกของความเป็นจริง
5. ใช้คนตรวจให้คะแนน โดยใช้ดุลพินิจประกอบการตัดสินใจ
6. ครูจำเป็นต้องรู้ทั้งวิธีการสอน และการประเมินรูปแบบใหม่

จากลักษณะรูปแบบการประเมินดังกล่าวข้างต้น ในประเทศสหรัฐอเมริกาได้ให้ความสนใจกับทางเลือกใหม่ของการประเมินผล ที่สามารถให้สารสนเทศเกี่ยวกับตัวผู้เรียนอย่างครอบคลุม

ที่สุด โดยเน้นความสามารถในการปฏิบัติในสถานการณ์ที่เป็นจริง จึงมีการนำเทคนิคการประเมินปฏิบัติการ (performance assessment) รูปแบบการประเมินนี้ เรียกกันในสหรัฐอเมริกาว่า เป็นทางเลือกใหม่ของการประเมิน (alternative assessment) หรือการประเมินความสามารถตามความเป็นจริง หรือการประเมินตามสภาพจริง (authentic assessment) (ภาวิณี ศรีสุขวัฒนานันท์, 2540) เป็นการประเมินที่ไม่เน้นการประเมินเฉพาะทักษะพื้นฐาน แต่จะเน้นการประเมินทักษะการคิดที่ซับซ้อนในการทำงานของผู้เรียน ความสามารถในการแก้ปัญหา และการแสดงออกที่เกิดจากการปฏิบัติในสภาพจริง ในการเรียนการสอนที่เน้นผู้เรียนเป็นศูนย์กลาง เป็นผู้ค้นพบ และผู้ผลิตความรู้ ผู้เรียนได้ฝึกปฏิบัติจริงรวมทั้งเน้นพัฒนาการเรียนรู้ของผู้เรียน เพื่อสนองจุดประสงค์ของหลักสูตรและความต้องการของสังคม การประเมินผลจากสภาพจริงจะแตกต่างจากการประเมินผลการเรียนหรือการประเมิน เพื่อรับรองผลงานแบบดั้งเดิมเพราะจะเน้นการให้ความสำคัญสำคัญกับพัฒนาการ และความต้องการช่วยเหลือการประสบความสำเร็จของผู้เรียนแต่ละคน มากกว่าการประเมินผลการเรียนที่มุ่งให้คะแนนผลผลิต และจัดลำดับที่ แล้วเปรียบเทียบกับกลุ่ม และจะแตกต่างจากการทดสอบทั่วไป เนื่องจากจะเป็นการวัดผลโดยตรง ในสภาพการแสดงออกจริงๆ ในเนื้อหาวิชา ซึ่งการทดสอบด้วยข้อสอบจะวัดได้เฉพาะความรู้ ทักษะบางส่วน และเป็นการวัดโดยอ้อมเท่านั้น (กรมวิชาการ, 2540)

การประเมินนี้จะมีลักษณะเด่นที่เน้นการประเมินพัฒนาการของนักเรียน และประสิทธิภาพการเรียนการสอน การประเมินดังกล่าวจะครอบคลุมสภาพความเป็นจริง และสอดคล้องกับการแสดงออกของนักเรียนทั้งกระบวนการ และผลผลิต ซึ่งเกี่ยวกับแฟ้มผลงาน (Portfolios) การบันทึกความเห็น แบบสำรวจรายงาน นิทรรศการ และโครงงาน เกี่ยวกับวิทยาศาสตร์ คณิตศาสตร์ และการเขียนทั้งวิชาต่างๆ

ในการปฏิบัติงานใดๆ ในชีวิตจริงนั้นมีแนวทางไปสู่ความสำเร็จหลายแนวทาง และงานหรือคำตอบที่ได้ก็มีใช้คำตอบที่ถูกเพียงคำตอบเดียวหรือมีผลอย่างใดอย่างหนึ่งเสมอไป ในการทำงานภาคปฏิบัติที่สอดคล้องกับสภาพชีวิตจริงของนักเรียนก็เช่นเดียวกัน งานหรือสถานการณ์ที่กำหนดให้นักเรียนทำ จะมีแนวทางไปสู่ความสำเร็จของงาน และวิธีการหาคำตอบหลายแนวทาง คำตอบที่ได้ก็อาจมีใช่เป็นไปตามแนวทางที่กำหนดไว้เสมอไป จึงทำให้การตรวจให้คะแนนไม่สามารถให้ได้อย่างชัดเจน แน่นนอนเหมือนการตรวจให้คะแนนข้อสอบแบบเลือกตอบ ซึ่งอาจตรวจให้คะแนนด้วยเครื่องตรวจกระดาษคำตอบก็ได้ ดังนั้นการตรวจให้คะแนนการปฏิบัติที่ตัดสินใจโดยมนุษย์ จึงต้องกำหนดแนวทางเป็นการให้คะแนนไว้อย่างชัดเจน

สิ่งที่ใช้เป็นแนวทางในการประเมินชิ้นงานคือ scoring rubric หรือเกณฑ์การให้คะแนน ซึ่งมีลักษณะเป็นระดับผลการดำเนินงาน โดยแต่ละระดับเป็นรายการที่บรรยายถึงคุณภาพของผลงานในประเด็นที่ต้องการทำการประเมินเป็นการตอบคำถามว่าผู้ถูกประเมินทำอะไรได้สำเร็จ

หรือว่ามีความสำเร็จในขั้นใด ผลงานเป็นอย่างไร Farr and Tone (1994) กล่าวว่า เกณฑ์การให้ระดับความสำเร็จของงานนี้โดยทั่วไปเสนอในรูปแบบตาราง 2 มิติ คือ มิติของระดับคุณภาพที่อาจจะแบ่งเป็น 3-6 ระดับก็ได้ และมิติของประเด็นที่นำมาประเมินหรือรายการประเมินนั่นเอง Hart (1994) เสนอว่าเกณฑ์การให้คะแนนผลงานมี 2 แบบ คือ เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) ซึ่งมีรายละเอียดดังนี้

1. เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) คือ เกณฑ์การให้คะแนนโดยพิจารณาความสำเร็จในภาพรวมของการปฏิบัติงาน ซึ่งอาจเป็นชิ้นงาน 1 ชิ้น หรือ ผลงาน 1 แฟ้ม โดยระดับความสำเร็จอาจแบ่งเป็น 4-6 ระดับก็ได้

2. เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) เพื่อให้เห็นคุณภาพของการปฏิบัติงานในรายละเอียดปลีกย่อย การประเมินจึงมีการวิเคราะห์แยกมิติหรือองค์ประกอบที่จะพิจารณาให้ระดับคะแนน ซึ่งแต่ละมิติหรือองค์ประกอบก็จะแบ่งระดับคุณภาพความสามารถไว้เช่นเดียวกับในเกณฑ์การประเมินภาพรวม (อมรรัตน์ พันธุ์งาม, 2543)

การประเมินผลงานเป็นกระบวนการที่มีความสำคัญมาก ซึ่งเป็นภาระที่ท้าทาย ผู้ที่ใช้ผลงานในการประเมินผลการเรียน คือจะมีวิธีการปฏิบัติอย่างไร ในการพิจารณาว่าผลงานมีคุณภาพหรือไม่ โดยทั่วไปจะต้องพิจารณาในเรื่องของความเที่ยง (reliability) และความตรง (validity) กล่าวคือผลการประเมินมีความคงที่หรือไม่ และคะแนนที่ได้มีความหมายว่าอย่างไร คะแนนที่ได้สามารถสรุปพาดพิงไปสู่การตัดสินใจที่ถูกต้องเกี่ยวกับนักเรียน วิชา หรือโรงเรียนได้หรือไม่ ในการศึกษาว่าผลงานนั้นมีคุณภาพด้านความเที่ยงหรือไม่ จะพิจารณาความสอดคล้องกันของผู้ประเมินกล่าวคือ ผู้ประเมินทั้งหลายซึ่งมีหน้าที่ตัดสินผลงานนักเรียน ต้องมีความสอดคล้องกันเกี่ยวกับคะแนนที่ให้แก่ผลงานของนักเรียน และสอดคล้องกันในเรื่องวิธีประเมิน การให้คะแนนผลงานของนักเรียน ของผู้ประเมินแต่ละคนในงานชิ้นเดียวกันควรใกล้เคียงกัน แต่ถ้าปราศจากซึ่งความสอดคล้องกันดังกล่าว แสดงว่าคะแนนที่ได้จากการประเมินจะขึ้นอยู่กับผู้ประเมินแต่ละคน ไม่ได้ขึ้นอยู่กับคุณภาพของชิ้นงานของนักเรียน (สมศักดิ์ ภูวิภาดาบรรณ, 2544)

การหาความเที่ยงของการประเมินการปฏิบัติงาน โดยวิธีการหาความสอดคล้องกันระหว่างผู้ประเมินเป็นที่ยอมรับในฐานะที่ใช้ตัดสินใจว่าผลงานนั้นมีคุณภาพหรือไม่ Herman and Winter (1994) พบว่าแหล่งสำคัญของคะแนนความเที่ยงขึ้นอยู่กับสิ่งต่อไปนี้ 1) ความคงที่ของคะแนน แม้ว่าจะวัดเมื่อใดก็ตาม 2) ความคงที่ของคะแนนที่เกิดจากผู้ประเมินแต่ละคน ในกลุ่มหรือคู่ของผู้ประเมิน และขึ้นอยู่กับ 3) ผลงานหรือบริบท (context) ที่ได้รับการประเมินด้วย

ส่วนมากวิธีการทั่วไปที่ใช้ในการประมาณค่าความเที่ยงระหว่างผู้ประเมินนั้น ได้มาจากการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิง (generalizability theory) อธิบายครั้งแรก โดย Cronbach et al. (1963) วิธีการนี้将有ความยืดหยุ่นที่จะรวมแหล่งความคลาดเคลื่อนของการวัดรวมทั้งผู้ประเมิน ผลงาน จำนวนครั้ง และรูปแบบของการประเมิน ค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) สามารถแปรความเป็นดัชนีความเห็นพ้องต้องกันของผู้ประเมินโดยจะต้องเป็นผู้ประเมินเดียวกัน ฝึกในทิศทางเดียวกัน และตรวจให้คะแนนภายใต้เงื่อนไขเดียวกัน ความสอดคล้องระหว่างผู้ประเมินเป็นประเด็นที่สำคัญมากสำหรับการประเมินการปฏิบัติ (performance assessment) ดังเช่น Korelz et al. (1994) แนะนำว่า ถึงแม้ว่าค่าความเที่ยงของผู้ประเมินจะมีขอบเขตในเรื่องของคะแนนจากทฤษฎีเทียบกับข้อเท็จจริงที่ได้จากการวัด ค่าความเที่ยงของผู้ประเมินที่สูงไม่ได้หมายความว่าคะแนนความเที่ยงจะเป็นที่น่าพอใจ ดังนั้นค่าความเที่ยงระหว่างผู้ประเมินจึงจำเป็น นอกจากนี้ ชัยพฤกษ์ เสรีรักษ์ และคณะ (2540) กล่าวว่า การวิเคราะห์ความเที่ยงของการวัดผลงานของนักเรียนอย่างต่อเนื่องนั้นจะวิเคราะห์ด้วยทฤษฎีความเที่ยงแบบดั้งเดิม เช่น การแบ่งครึ่งข้อสอบแบบสัมประสิทธิ์แอลฟาของครอนบาค หรือใช้สูตรของ คูเดอร์ ริชาร์ดสัน ย่อมไม่เหมาะสมเพราะทฤษฎีเหล่านี้มีข้อตกลงว่า ข้อสอบวัดสิ่งเดียวกันมีลักษณะคู่ขนาน และดำเนินการสอบครั้งเดียวแต่ผลงานที่ผู้เรียนผลิตออกมาระหว่างการเรียนนั้นมีมากมาย และมีลักษณะแตกต่างกัน ในรูปของการบ้าน งานโครงงาน ผลการทำงานเป็นกลุ่ม ตลอดจนพฤติกรรมที่ครูสังเกต เช่น ความรับผิดชอบ ความมีวินัย ความซื่อสัตย์ การวัดผลตามลักษณะดังกล่าว จากทฤษฎีความเที่ยงที่เรียกว่า ทฤษฎีการสรุปอ้างอิงเป็นทฤษฎีที่เหมาะสมในการวิเคราะห์ค่าความเที่ยงของการวัดผลงานที่ผลิตออก

การศึกษางานวิจัยที่เกี่ยวข้องกับการประเมินการปฏิบัติงานนั้นยังมีไม่มากเท่าที่ควร จึงทำให้ผู้วิจัยสนใจศึกษาเปรียบเทียบค่าความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนที่ได้จากการประเมินผลงานการเขียนจากผู้ประเมิน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และผู้ประเมินอีก 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) ว่าจะมีค่าความเที่ยงแตกต่างกันอย่างไร และศึกษาว่าจำเป็นต้องใช้จำนวนผู้ประเมินและจำนวนงานเขียนเท่าใดที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนต่างกัน เกณฑ์การให้คะแนนใดที่จะมีค่าความเที่ยงสูง สามารถนำมาใช้แล้วทำให้การวัดมีความแม่นยำมากที่สุด

คำถามวิจัย

จากการศึกษาประเด็นเกี่ยวกับเกณฑ์การให้คะแนน ซึ่งมี 2 รูปแบบ ได้แก่ เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) เมื่อนำมาใช้ในการประเมินผลงานการเขียนทำให้เกิดคำถามวิจัยดังนี้

1. ค่าความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงที่ได้จากการให้คะแนนงานเขียนของผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) จะมีค่าความเที่ยงแตกต่างกันอย่างไร
2. จำนวนผู้ประเมินและจำนวนงานเขียนเท่าใดที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)

วัตถุประสงค์ของการวิจัย

1. เพื่อเปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนจากผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)
2. เพื่อหาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)

ขอบเขตของการวิจัย

1. เนื่องจากประเมินครั้งนี้มีขั้นตอนการเก็บรวบรวมข้อมูลที่ต้องอาศัยความร่วมมือในการวิจัย คือ ให้ครูช่วยประเมินผลงานการเขียน ดังนั้นผู้วิจัยจึงเลือกโรงเรียนที่ยินดีให้ความร่วมมือ ในเขตกรุงเทพมหานคร จำนวน 1 โรงเรียน

2. ประชากรและกลุ่มตัวอย่าง

2.1 ประชากรที่ใช้ในการศึกษาคั้งนี้เป็น ผู้ประเมินซึ่งเป็นครูผู้สอนวิชาภาษาไทย ชั้นประถมศึกษาปีที่ 5-6 ในโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาเอกชน ในเขต กรุงเทพมหานครและปริมณฑล และผลงานการเขียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนที่ผู้วิจัยคัดเลือก ในเขตกรุงเทพมหานคร จำนวน 1 โรงเรียน ได้แก่ โรงเรียนไผทอุดมศึกษา

2.2 ตัวแปรที่ใช้ในการวิจัย

2.2.1 ตัวแปรอิสระ (independent variable) ประกอบด้วย

- 1) จำนวนผู้ประเมิน
- 2) จำนวนงานเขียน
- 3) เกณฑ์การให้คะแนน

2.2.2 ตัวแปรตาม (dependent variable) ได้แก่ ค่าความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient)

คำจำกัดความที่ใช้ในการวิจัย

ผลงานการเขียน หมายถึง ผลงานที่แสดงความสามารถและความก้าวหน้าทางด้านการเขียนในแฟ้มสะสมงานวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2544 ประกอบด้วย การเขียนเรื่องประกอบภาพ การเขียนจดหมาย และการบันทึกความรู้เรื่องที่อ่าน

เกณฑ์การให้คะแนนแบบภาพรวม หมายถึง แนวทางการให้คะแนนผลงานการเขียนวิชาภาษาไทยโดยพิจารณาจากภาพรวมของผลงานการเขียนว่ามีคุณลักษณะอย่างไรและมีการอธิบายคุณภาพของงานเป็นระดับจาก 1 ถึง 4

เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ หมายถึง แนวทางการให้คะแนนผลงานการเขียนวิชาภาษาไทยโดยให้คะแนนแบบแยกองค์ประกอบ เช่น ด้านเนื้อหา การลำดับเรื่อง การใช้ภาษา เป็นต้น และมีการอธิบายคุณภาพของงานแต่ละรายการเป็นระดับจาก 1 ถึง 4

ผู้ประเมิน หมายถึง ครูผู้สอนวิชาภาษาไทยชั้นประถมศึกษาปีที่ 5-6 มีประสบการณ์ด้านการสอนวิชาภาษาไทยไม่น้อยกว่า 5 ปี

สัมประสิทธิ์การสรุปอ้างอิง หมายถึง คุณภาพด้านความเที่ยงของคะแนนที่ได้จากการประเมินงานเขียนโดยใช้ทฤษฎีการสรุปอ้างอิง (Generalizability Theory) ซึ่งคำนวณจากการวิเคราะห์ความแปรปรวนโดยใช้โปรแกรมสำเร็จรูป GENOVA

ข้อตกลงเบื้องต้น

ผู้วิจัยได้กำหนดเกณฑ์คุณสมบัติผู้ประเมิน คือ ต้องเป็นครูที่มีประสบการณ์ในการสอนวิชาภาษาไทยไม่น้อยกว่า 5 ปี และผู้ประเมินที่เข้าร่วมในการวิจัยได้ประเมินผลงาน ด้วยความตั้งใจและเต็มใจ

ประโยชน์ที่ได้รับจากการวิจัย

1. ทำให้ทราบจำนวนผู้ประเมิน จำนวนผลงานที่เหมาะสมสำหรับการประเมินงานเขียน เมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน
2. นักวัดผล และครูสามารถพิจารณาเลือกใช้เกณฑ์การให้คะแนนผลงาน จำนวนผู้ประเมิน และจำนวนผลงานที่เหมาะสม สามารถนำไปปฏิบัติได้



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยนำเสนอ แบ่งออกเป็น 4 ตอน คือ

ตอนที่ 1 การประเมินแบบดั้งเดิมและการประเมินตามสภาพจริง

ตอนที่ 2 เกณฑ์การให้คะแนน

ตอนที่ 3 ทฤษฎีการสรุปอ้างอิง

ตอนที่ 4 งานวิจัยอื่นๆ ที่เกี่ยวข้องกับเกณฑ์การให้คะแนน และทฤษฎีการสรุปอ้างอิง

ตอนที่ 1 การประเมินแบบดั้งเดิมและการประเมินตามสภาพจริง

การประเมินแบบดั้งเดิม (traditional assessment)

การประเมินแบบดั้งเดิม (traditional assessment) เป็นการประเมินผลงานแบบประเพณีนิยมที่ได้ปฏิบัติกันมาจนถึงทุกวันนี้ การประเมินแบบดั้งเดิมนี้ ผู้เรียน และนักการศึกษา มักมุ่งเน้นไปที่คะแนนจากการทดสอบมากกว่าการใช้การทดสอบเป็นเครื่องมือประเมินการเรียนรู้ ที่พบบ่อยก็คือผู้สอนมักจะสอนข้อสอบให้แก่ผู้เรียน หรือติวข้อสอบให้แก่ผู้เรียนเพื่อให้ผู้เรียนทำข้อสอบให้ผ่านให้ได้ โดยปราศจากการเน้นเรื่องการพัฒนาทักษะต่างๆ (Kellaghan and Madaus, 1991; Worthen, 1998; Michell, 1992) นอกจากนี้การประเมินผลแบบดั้งเดิมนิยมใช้แบบทดสอบชนิดเลือกตอบ เน้นการจัดอันดับผู้เรียน เน้นพุทธิพิสัย เน้นทักษะตามลำดับขั้น เน้นความเป็นเอกัตบุคคล ความรู้ในเรื่องใดจะมีความหมายเดียวกันในทุกที่ แยกกระบวนการออกจากผลผลิต ผู้เรียนเปรียบเหมือนผู้รับ ผู้สอนมีอำนาจตัดสินใจโดยผู้เรียนมักไม่มีส่วนร่วม และสิ่งที่ประเมินต้องเป็นปรนัยแยกข้อเท็จจริงออกจากค่านิยม (สมศักดิ์ ภาวธาวรรณ, 2544) จึงพบว่ากระบวนการทั้งหมดยังมีข้อจำกัดในการใช้เพื่อปรับปรุงแก้ไขจุดอ่อนในด้านต่างๆของผู้เรียนเป็นรายบุคคล บ่อยครั้งพบว่า ผู้เรียนไม่สามารถประมาณความสามารถของตน และไม่รู้จักจุดอ่อนที่ควรปรับปรุงแก้ไข เพราะไม่ทราบว่ามาตรฐานของการปฏิบัติที่ควรจะเป็นตามความคาดหวังของหลักสูตร หรือของผู้สอนเป็นอย่างไร ผู้สอนก็ขาดทิศทางของการจัดกิจกรรมการสอนที่ให้โอกาสนักเรียนได้พัฒนาตนเองอย่างเต็มที่ (ภาวธาวรรณ ศรีสวัสดิ์, 2540)

การประเมินตามสภาพจริง (authentic assessment)

ความหมายและแนวคิดของการประเมินตามสภาพจริง

ความหมายของการประเมินตามสภาพจริงมีผู้ให้ความหมายไว้มากมายตามแต่ใครจะเน้นไปทางใด แต่โดยสรุปแล้วจะเน้นที่ความจริง (realistic) ในชีวิตจริง โดยมีผู้ให้ความหมายต่างกันไป ดังนี้

Ronis (2000) ได้ให้ความหมายของการประเมินตามสภาพจริงว่า เป็นการประเมินโดยใช้ชิ้นงาน (tasks) ของนักเรียน ซึ่งสามารถให้ความหมายโดยตรง

Nitko (1996) กล่าวว่า การประเมินตามสภาพจริง เป็นการประเมินผลงานของนักเรียนทั้งหมดที่มีความคุ้มค่า (worthwhile) ความสำคัญ (significant) และความหมาย (meaningful) โดยเปิดโอกาส และอนุญาตให้นักเรียนแสดงความสามารถได้ในหลายแนวทาง

Cole et al. (2000) กล่าวว่า การประเมินตามสภาพจริง หมายถึง เป็นการประเมินที่ใช้หลักฐานจริง การปฏิบัติที่สมบูรณ่เกี่ยวกับการอ่าน การเขียน การค้นหา การแก้ปัญหา ความคิดสร้างสรรค์ และการพูด มีการสังเกตเกี่ยวกับงานของนักเรียน กระบวนการประเมินต้องอาศัยทั้งนักเรียน และครู การประเมินนี้นักเรียนต้องสามารถนำไปใช้ในชีวิตจริงได้

กรมวิชาการ (2540) กล่าวว่า การประเมินตามสภาพจริงหมายถึง กระบวนการสังเกต การบันทึก การรวบรวมข้อมูลจากงาน และวิธีการที่นักเรียนทำ เพื่อเป็นพื้นฐานของการตัดสินใจในการศึกษาถึงผลกระทบต่อเด็กเหล่านั้น ไม่เน้นการประเมินเฉพาะทักษะพื้นฐาน แต่เป็นการเน้นการประเมินทักษะการคิดที่ซับซ้อน ในการทำงานของนักเรียน ความสามารถในการแก้ปัญหา และการแสดงออกที่เกิดจากการปฏิบัติในสภาพจริง

สำนักงานการประถมศึกษาแห่งชาติ (2540) กล่าวถึงความหมายของ การประเมินผลจากสภาพจริงว่าหมายถึง การประเมินผลโดยเน้นการปฏิบัติจริง เป็นการประเมินการกระทำ การแสดงออกในหลายๆด้านของนักเรียน ตามสภาพความเป็นจริงทั้งในและนอกชั้นเรียน หรือสถานที่อื่น ๆ นอกโรงเรียน

ส่วน ส. วาสนา ประवालพฤษ์ (2539) ให้ความหมายว่า การประเมินตามสภาพจริง เป็นการวัดโดยเน้นให้ผู้เรียนได้นำความรู้ แนวคิดในวิชาต่างๆที่เรียนเพื่อนำมาแก้ปัญหาโดยใช้ทักษะการคิดที่ซับซ้อนมากกว่าที่จะถามความสามารถขั้นต้น หรือความสามารถย่อยๆเป็นการวัดผู้เรียนโดยรวมทั้งด้านความคิด เจตคติ และการกระทำไปพร้อมๆกัน

จากความหมายของการประเมินตามสภาพจริงที่กล่าวมาข้างต้น กล่าวโดยสรุปได้ว่าการประเมินตามสภาพจริง หมายถึง การประเมินผลงานของนักเรียนที่มีการปฏิบัติ การแสดงความสามารถของนักเรียนโดยใช้ทักษะการคิดที่ซับซ้อน เพื่อนำไปใช้แก้ปัญหาสถานการณ์ที่เป็นจริงในทุกบริบทเท่าที่จะเป็นไปได้

แนวคิดการประเมินตามสภาพจริง

จากการประเมินผลแนวทางเดิมที่โดยมาก จะเป็นข้อสอบแบบดั้งเดิมที่สอบวัดด้วยข้อสอบมาตรฐานที่โดยมาก มักจะเป็นข้อสอบแบบเลือกตอบถูกผิด และเติมคำสั้นๆที่ใช้ในการตอบคำถามเพื่อการทดสอบจะไม่สามารถวัดความก้าวหน้า และผลผลิตที่ซับซ้อนที่เกิดขึ้นกับผู้เรียนหลายๆด้าน ดังผลที่ได้จากการทำโครงการ ซึ่งการประเมินแบบดั้งเดิมก่อให้เกิดความเฉื่อยชา แยกเนื้อหาความรู้เป็นส่วนย่อยๆ ส่วนมากเน้นความจำมากกว่าความคิดในขั้นสูง และการนำไปใช้ จึงเกิดความคิดใหม่ๆ ที่มีความจำเป็นต่อผสมผสานหลักสูตร การสอน และการประเมินผลเพื่อให้สามารถกล่าวได้อย่างเต็มที่ถึงผลการเรียนรู้ที่เกิดขึ้นในห้องเรียนที่มีความหมาย รวมทั้งผลสัมฤทธิ์ที่เกิดขึ้นแก่นักเรียนในสภาพที่ตรงกับความเป็นจริง

การประเมินผลจากสภาพจริงจะแตกต่างจากการประเมินผลการเรียน หรือการประเมินเพื่อรับรองผลงานแบบดั้งเดิมเพราะจะเน้นการให้ความสำคัญกับการพัฒนาการ ความต้องการช่วยเหลือ และประสบความสำเร็จของผู้เรียนแต่ละคนมากกว่าการประเมินผลการเรียนที่มุ่งให้คะแนนผลผลิต และจัดลำดับที่แล้วเปรียบเทียบกับกลุ่ม และจะแตกต่างจากการทดสอบทั่วไปเนื่องจากจะเป็นการวัดผลโดยตรง ในสภาพการแสดงออกจริงๆ ในเนื้อหาวิชา ซึ่งการทดสอบด้วยข้อสอบจะวัดผลได้เฉพาะความรู้ และทักษะบางส่วนและเป็นการวัดโดยอ้อมเท่านั้น นอกจากนี้การประเมินจากสภาพจริงจะมีความต่อเนื่องในการให้ข้อมูลเชิงคุณภาพที่เป็นประโยชน์ต่อครูผู้สอน ได้ใช้เป็นแนวทางการจัดกิจกรรมการสอนให้เหมาะสมกับแต่ละบุคคลได้ (กรมวิชาการ, 2540)

การประเมินตามสภาพจริงจะต้องมีการดำเนินการอย่างต่อเนื่องตลอดเวลา และมีความสัมพันธ์ซึ่งกันและกันโดยรวมทั้งการเรียนรู้ การสอน และการประเมินเข้าด้วยกัน (Boron, 1991 อ้างถึงใน Nitko, 2000) กล่าวว่า การประเมินตามสภาพจริงจะต้องมีลักษณะดังนี้คือ

1. เน้นการนำไปใช้ นักเรียนสามารถที่จะทำอะไรที่จะประเมินได้ และต้องเรียนรู้อะไรบ้าง
2. เน้นการประเมินทางตรง มีการประเมินเป้าหมายของการเรียนรู้ทั้งทางตรงและทางอ้อม
 - 2.1 ใช้ปัญหาที่แท้จริง กรอบของชิ้นงาน จะต้องมีความที่เป็นจริงสูง นักเรียนควรที่จะสามารถจดจำ แก้ปัญหาได้
 - 2.2 สนับสนุนการคิดที่กว้าง กรอบของชิ้นงานจะสนับสนุนให้มีคำตอบถูกมากกว่า 1 คำตอบโดยนักเรียนทำงานร่วมกัน

นอกจากนี้ Cole et al. (2000) กล่าวว่า การประเมินตามสภาพจริงต้องมีลักษณะดังนี้

1. ช่วยเหลือในการเรียนรู้แต่ละบุคคล
2. สนับสนุนการสอนที่ดี

3. เกี่ยวข้องกับผลลัพธ์ของหลักสูตร
4. สนับสนุนการเรียนรู้ในระดับสูง
5. การพัฒนาตามทัศนวิสัย
6. ใช้การทดสอบน้อยลง
7. ใช้เวลาอย่างมีประสิทธิภาพ
8. รายงานต้องให้ข้อมูลที่มีความหมาย
9. สนับสนุนการมีส่วนร่วมของ ผู้ปกครอง ครู และนักเรียน
10. นักเรียนมีการสะท้อนข้อมูลย้อนกลับ

ในกระบวนการจัดการเรียนการสอนโดยเน้นสภาพจริงของผู้เรียนนั้น การวัดและประเมินผลตามสภาพจริงของผู้เรียน มีความสำคัญต่อการพัฒนากระบวนการเรียนการสอนมาก ซึ่งการประเมินตามสภาพจริงมีความสำคัญดังนี้ (กรมวิชาการ, 2540)

1. การเรียนการสอน และการวัดผลประเมินตามสภาพจริง จะเอื้อให้นักเรียนสามารถเรียนรู้ได้อย่างเต็มศักยภาพของแต่ละบุคคล เพราะเน้นลักษณะสำคัญดังนี้

- 1.1 เน้นให้นักเรียนได้แสดงออก/สร้างสรรค์/ผลิตหรือทำงาน

- 1.2 ดึงเอาความคิดชั้นสูง ความคิดซับซ้อน และทักษะการแก้ปัญหาออกมาได้

- 1.3 ผลสัมฤทธิ์ของนักเรียนเป็นผลมาจากการเรียนการสอนที่สอดคล้องกับความเป็นจริง ชีวิตประจำวัน

- 1.4 กระตุ้นให้เกิดการประยุกต์สู่โลกของความเป็นจริง

2. การประเมินตามสภาพจริง จะเอื้อต่อการเรียนการสอนที่เน้นนักเรียนเป็นศูนย์กลางมากกว่าการเรียนการสอนที่เกิดจากครูเป็นผู้บอกความรู้ โดยครูจะเป็นเพียงผู้ชี้แนะว่าควรจัดเนื้อหาสาระอย่างไร นักเรียนจะเรียนรู้จากการกระทำมากขึ้น มีความสนใจในบทเรียนมากขึ้น การบ่งชี้ความสามารถที่แท้จริงของผู้เรียน มิใช่เป็นเพียงทำข้อสอบได้คะแนนสูงเท่านั้น การประเมินผลจากสภาพจริงจะแสดงให้เห็นว่านักเรียนทำอะไรได้มากกว่าจะบอกว่าคุณเรียนรู้อะไร

3. สังคมมนุษย์ในปัจจุบันและอนาคต จะเป็นสังคมที่รวมกันเป็นกลุ่มใหญ่เป็นสังคมโลก เนื่องจากความเจริญด้านเทคโนโลยีสารสนเทศ คนในสังคมจะมีการแข่งขันกันมากขึ้น วิถีชีวิตของคนในสังคมมีความซับซ้อนมากขึ้น การจัดการศึกษาแบบให้ผู้เรียนแยกเป็นส่วนๆโดยการทำแบบฝึกหัดจากสมุดแบบฝึกหัดหรือใบงาน แล้วตอบคำถามไม่น่าจะพอเพียงสำหรับการเตรียมเยาวชนให้ดำรงชีวิตในสังคมอย่างปกติสุข ดังนั้นการให้ผู้เรียนได้สร้างงานเพื่อแสดงให้เห็นถึงความสามารถในการพัฒนาการเรียนรู้ และการบูรณาการวิชาต่างๆเข้าด้วยกันเป็นโครงการภาระงาน จึงจำเป็นอย่างยิ่งในการให้การศึกษแก่เยาวชนของชาติในปัจจุบัน ซึ่งครูจะต้องแสดง

ความรับผิดชอบ (accountability) ต่อสังคมด้วยการปรับเปลี่ยนวิธีการสอนและการประเมินผล ให้สามารถตอบสนองความต้องการของสังคมด้วย ผลงานของนักเรียนที่ปรากฏจะเป็นหลักฐานที่สำคัญที่แสดงถึงความรับผิดชอบของครูผู้สอน

4. โดยทั่วไปครูมักจะมองภาพการสอน การเรียนรู้ของนักเรียน และการประเมินผลงานเป็นงานที่แยกออกจากกัน โดยครูให้ความรู้ข้อมูลต่างๆพอนักเรียนเกิดการเรียนรู้แล้วจึงทำการประเมินผล ซึ่งใช้วิธีเรียกว่าการสอบทำให้นักเรียนมีความวิตกกังวล ไม่มีความสุข ในการเรียน เพราะการสอบจะเป็นการเน้นการจับผิด หากจุดด้อยของผู้เรียนในขณะที่เจตนาที่แท้จริงของการประเมินผลคือ ช่วยพัฒนาการเรียนรู้ของนักเรียนและการสอนของครูเป็นการค้นหา จุดดีของนักเรียนเพื่อเสริมสร้างผู้เรียนให้พัฒนาเต็มศักยภาพ ดังนั้นการประเมินผลการเรียนรู้ การสอน จึงมีความสัมพันธ์กันอย่างใกล้ชิด และเกิดขึ้นในเวลาเดียวกัน

5. ในการเรียนเพื่อรอบรู้นั้นเมื่อนักเรียนไม่ผ่านจุดประสงค์ย่อยๆ ครูก็ซ่อมเสริมด้วยวิธีการต่างๆ ซึ่งเป็นภาระแก่ผู้สอนปัญหานี้จะหมดไป ถ้าครูจัดการเรียนการสอนและประเมินผลจากสภาพจริงเพราะมีรายงาน หรือหลักฐานการปฏิบัติงานของนักเรียนขณะเรียน และมีการพัฒนา แก้ไขข้อบกพร่องมาโดยตลอด รวมทั้งมีการบันทึก การปฏิบัติงาน การรายงาน การแสดงความคิดเห็นต่างๆของผู้เรียนแต่ละคน ช่วยให้ผู้เรียนได้เรียนรู้อย่างเต็มที่รวมทั้งมีข้อมูลยืนยันความสามารถที่แท้จริงของผู้เรียน เพื่อแสดงว่าผู้เรียนสามารถทำอะไรได้มากกว่าจะบอกว่าเขาทำอะไร แค่ไหน

กล่าวโดยสรุป แนวคิดหลักของการประเมินตามสภาพจริง เป็นการประเมินผลการเรียนจากการที่ครูเฝ้าดูนักเรียนโดยอาศัยการสังเกต และเก็บข้อมูลอย่างต่อเนื่อง ตลอดระยะเวลาที่ผู้เรียนอยู่กับครูเป็นกิจกรรมที่สอดแทรกในสภาพการเรียนการสอนในชีวิตประจำวัน การจัดกิจกรรมการเรียนการสอนจะต้องเน้นให้ผู้เรียนเป็นศูนย์กลาง เป็นผู้ค้นพบความรู้ เป็นผู้สร้างงาน นักเรียนมีโอกาสได้ฝึกปฏิบัติจริงๆ การประเมินตามสภาพจริงจะครอบคลุมไปถึงการประเมินจากแฟ้มสะสมงาน การจัดนิทรรศการ การทำงานเป็นกลุ่ม การแสดงการกระทำ การแสดงออก โครงการ การนำเสนอผลงาน แบบฝึกหัด บันทึกรายวัน

ตอนที่ 2 เกณฑ์การให้คะแนน

เกณฑ์การให้คะแนนผลงาน (scoring rubric)

การประเมินตามสภาพจริงมักนิยมใช้เกณฑ์การให้คะแนนรูบรีคเพื่ออธิบายความสัมฤทธิ์ผลของผู้เรียน มีนักวิชาการหลายท่านได้ให้ความหมายของเกณฑ์การให้คะแนน (scoring rubric) ดังนี้

ชัยฤทธิ์ ศิลาเดช (2540) กล่าวว่า เกณฑ์การให้คะแนนผลงาน เป็นเครื่องมือที่ใช้เป็นแนวทางในการประเมินผลงานของนักเรียน เรียกว่า “รูบรีค” (rubric มาจากภาษาลาตินว่า rubricaterra) เป็นคำที่ใช้ในสมัยโบราณเกี่ยวกับศาสนา หมายถึงการทำเครื่องหมายสีแดงไว้บนสิ่งที่สำคัญ ดังนั้น รูบรีคก็คือ แนวทางการให้คะแนนซึ่งต้องกำหนดมาตรวัด (scale) และรายการของคุณลักษณะที่บรรยายถึงความสามารถในการแสดงออก หรือคุณลักษณะแต่ละจุดในมาตรวัดไว้อย่างชัดเจน

สมศักดิ์ ภูวิภาดาวรรณ (2544) ได้ให้ความหมายของรูบรีค ดังนี้ รูบรีค คือ เครื่องมือในการให้คะแนน (scoring tool) ที่มีการระบุเกณฑ์ (criteria) ประเมินชิ้นงาน และคุณภาพ (quality) ของชิ้นงานในแต่ละเกณฑ์ ซึ่งการสร้างรูบรีคอาจทำขึ้นเพื่อจุดประสงค์ดังนี้ 1) เพื่อการประเมินกระบวนการ (process) เช่น ประเมินการเรียนรู้เป็นทีม กลยุทธ์การสัมภาษณ์ เป็นต้น 2) เพื่อประเมินผลผลิต (product) เช่น การประเมินแฟ้มสะสมงาน รายงานการวิจัย นิทรรศการ เป็นต้น 3) เพื่อการประเมินปฏิบัติ (performance) เช่น การประเมินการนำเสนอปากเปล่า การอภิปราย การสาธิต เป็นต้น

Goodrich (1997) กล่าวว่า รูบรีค (rubric) เป็นเครื่องมือการให้คะแนน ซึ่งจะแสดงเกณฑ์สำหรับชิ้นงานจะเป็นการเชื่อมต่อกัน ในการแบ่งลำดับของคุณลักษณะแต่ละเกณฑ์ จากดีมาจนถึงต้องปรับปรุงแก้ไข ประโยชน์ของรูบรีคมีดังนี้คือ

1. รูบรีคสามารถจะพัฒนาการปฏิบัติงานของนักเรียนคล้ายกับเป็นการกำกับติดตาม โดยทำให้ความคาดหวังของครูมีความชัดเจน และเป็นการแสดงว่านักเรียนจะพบกับสิ่งที่คาดหวังได้อย่างไร ผลที่ได้มีจะเป็นการบอกปรับปรุงแก้ไขในคุณลักษณะของงานนักเรียน และในการเรียนรู้
2. รูบรีคสามารถช่วยตัดสินใจเกี่ยวกับคุณลักษณะงานของนักเรียนด้วยตนเอง และงานอื่นๆ เมื่อรูบรีคนำมาใช้ในการประเมินตนเองและกลุ่ม จะทำให้เพิ่มความสามารถแก้ปัญหาในงานนั้น
3. รูบรีคจะช่วยลดเวลาของครูที่ใช้ในการประเมินงานของนักเรียนได้
4. รูบรีคใช้ง่าย และมีการอธิบายที่ชัดเจน

Ronis (2000) กล่าวว่า ฐุบรึคเป็นแนวทางการให้คะแนนผลงาน หรือการปฏิบัติของนักเรียน ฐุบรึคจะเป็นตัวแทนคุณลักษณะที่สังเกตได้จากผลงาน กระบวนการ หรือการปฏิบัติ ฐุบรึคจะช่วยให้ระบบการให้คะแนนมีความง่ายในการเรียนรู้และนำไปใช้ จะช่วยให้นักเรียนประเมินตามระดับผลสัมฤทธิ์และตัวชี้วัดที่จะทำให้มีการพัฒนา และปรับปรุงผลงานการปฏิบัติ ฐุบรึคจะเป็นส่วนสำคัญที่จะช่วยให้ครู สามารถตอบคำถามดังนี้

1. เป้าหมายการปฏิบัติของนักเรียนคืออะไร
2. อะไรเป็นสิ่งที่นักเรียนต้องรู้อย่างแท้จริงโดยขึ้นอยู่กับเกณฑ์มาตรฐาน
3. อะไรจะเป็นหลักฐานประกอบความเข้าใจของนักเรียน
4. ผลงานจะต้องมีลักษณะอย่างไรบ้าง
5. อะไรเป็นสิ่งที่ผิดพลาดแสดงถึงคะแนนที่ลดน้อยลง

จากที่กล่าวมาข้างต้นสรุปได้ว่าประโยชน์ของฐุบรึคหลายประการดังนี้

ช่วยพัฒนาผลงานของนักเรียน ขณะเดียวกันก็เป็นเครื่องมือในการติดตามพัฒนา ผลงานของนักเรียนเองด้วย เพราะฐุบรึคบอกไว้ชัดเจนว่า ครูคาดหวังอะไร และนักเรียนจะรู้ว่า จะก้าวไปถึงความคาดหวังนั้นได้อย่างไร

ช่วยให้นักเรียนหัดใช้ความคิดในการพิจารณาคุณภาพงานของตนเอง และผู้อื่น จากการใช้ฐุบรึคประเมินผลงานของตนเอง และเพื่อน การฝึกให้ทำงานหลายๆครั้งจะช่วยให้นักเรียน เป็นคนที่มีความรับผิดชอบต่องานของตนเอง

ช่วยลดเวลาที่ครูต้องใช้สำหรับการประเมินผลงานนักเรียน เพราะหลังจากนักเรียน ประเมินตนเอง และให้นักเรียนประเมินโดยใช้ฐุบรึคแล้ว ครูมักพบว่าสิ่งที่ครูต้องปรับปรุงมีไม่มาก

ครูสามารถปรับฐุบรึคให้เหมาะสมกับการประเมินผลงานนักเรียนกลุ่มต่างๆ ที่ต่างกัน

เป็นสิ่งที่ง่ายและอธิบายให้คนอื่นเข้าใจได้ง่าย เช่น อธิบายให้ผู้ปกครองทราบ

การกำหนดเกณฑ์การให้คะแนนของฐุบรึค (scoring rubric) ก็คือการตอบคำถามว่า นักเรียนทำอะไรได้สำเร็จ หรือว่ามีระดับความสำเร็จในขั้นต่างๆ เท่าใด หรือผลงานเป็นอย่างไร มีคุณภาพอยู่ในระดับใดนั่นเอง มีวิธีการกำหนด 2 แบบ คือ (กรมวิชาการ, 2540)

1. เกณฑ์การให้คะแนนเป็นภาพรวม หรือเกณฑ์รวม (holistic scoring rubric) คือ เกณฑ์การให้คะแนนชิ้นงานใดชิ้นงานหนึ่ง โดยคุณภาพรวมของชิ้นงานว่ามีคุณลักษณะเป็น อย่งไร แล้วเขียนอธิบายคุณภาพของงานหรือความสำเร็จของงานเป็นขั้นๆ โดยอาจแบ่งระดับ คุณภาพ ตั้งแต่ 0 – 4 หรือ 0 – 6 โดยมีวิธีการในการกำหนดคะแนนตามคุณภาพของงาน ดังนี้

1.1 แบ่งงานตามคุณภาพ เป็น 3 กอง คือ

กองที่ 1 ได้แก่ งานที่มีคุณภาพเป็นพิเศษ เขียนอธิบายลักษณะของงานที่มีคุณภาพเป็น
พิเศษ

กองที่ 2 ได้แก่ งานที่ยอมรับได้ เขียนอธิบายลักษณะของงานที่ยอมรับได้

กองที่ 3 ได้แก่ งานที่ยอมรับได้น้อย หรือ ยอมรับไม่ได้ เขียนอธิบายลักษณะของงาน
ที่ยอมรับได้น้อย จากนั้นก็นำงานแต่ละกองมาให้คะแนน เช่น กองที่ 1 จะให้คะแนนสูงสุดลดหลั่น
มาจนถึงกองที่ 3 ส่วนงานที่แสดงว่าไม่ได้ใช้ความพยายามเลยให้คะแนนเป็น 0

1.2 กำหนดระดับความผิดพลาด โดยพิจารณาจากความบกพร่องของคำตอบว่า
มีมากน้อยเพียงใดแล้วหักจากระดับคะแนนสูงสุดมาที่ระดับ เช่น การกำหนดระดับ คุณภาพ
ของงาน แก้ปัญหาโจทย์คณิตศาสตร์ อาจกำหนดดังนี้

คะแนน 4 หมายถึง คำตอบถูก แสดงเหตุผลถูกต้อง แนวคิดชัดเจน

คะแนน 3 หมายถึง คำตอบถูก เหตุผลถูกต้อง มีข้อผิดพลาดเล็กน้อย

คะแนน 2 หมายถึง เหตุผลหรือการคำนวณผิดพลาด แต่มีแนวทางที่จะนำไป

สู่คำตอบ

คะแนน 1 หมายถึง แสดงวิธีคิดเล็กน้อย แต่ไม่ได้คำตอบ

คะแนน 0 หมายถึง ไม่ตอบ หรือตอบไม่ถูก

1.3 อธิบายคุณภาพตามระดับจากสูงสุดถึงต่ำสุด เช่น เกณฑ์การให้คะแนน
ความสามารถในการทำความเข้าใจเนื้อหาสาระ อาจกำหนดดังนี้

คะแนน 4 หมายถึง การสาธิตหรือแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์
ครบถ้วนถูกต้องแม่นยำในหลักการความคิดรวบยอด ข้อเท็จจริงของงาน หรือสถานการณ์
ที่กำหนด รวมทั้งเสนอแนวคิดใหม่ que แสดงถึงความเข้าใจอย่างลึกซึ้งถึงกฎเกณฑ์ หรือลักษณะของ
ข้อมูล

คะแนน 3 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์ ครบถ้วน
ถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงาน หรือสถานการณ์ที่กำหนด

คะแนน 2 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจไม่สมบูรณ์ ครบถ้วน
ถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงาน หรือสถานการณ์ที่กำหนด

คะแนน 1 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจในหลักการความคิด
รวบยอด ข้อเท็จจริง หรือสถานการณ์ที่กำหนดน้อยมาก และเข้าใจไม่ถูกต้องในบางส่วน

คะแนน 0 หมายถึง ไม่แสดงความคิดเห็นใด ๆ

สิ่งที่ต้องกำหนดต่อไปนี้คือ การพิจารณาระดับคุณภาพใดจะเป็นที่ยอมรับได้ จะเห็นว่า
จากตัวอย่างที่ผ่านมา ตั้งแต่ระดับ 3 ขึ้นไปมีคำอธิบายถึงการแสดงออกที่ยอมรับได้ เพราะนักเรียน

แสดงออกถึงความเข้าใจที่สมบูรณ์ ครบถ้วน ถูกต้องในหลักการความคิดรวบยอด ข้อเท็จจริงของงาน หรือสถานการณ์ที่กำหนด

2. เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ หรือเกณฑ์ย่อย (analytic scoring rubric) เพื่อให้การมองคุณภาพของงานหรือความสามารถของนักเรียนได้อย่างชัดเจน จึงได้มีการแยกองค์ประกอบของการให้คะแนนเป็นหลายองค์ประกอบ หรือหลายรายการ และอธิบายคุณภาพของงานในแต่ละองค์ประกอบเป็นระดับ ซึ่งอาจกล่าวสรุปได้ว่าเกณฑ์การให้คะแนนแบบนี้ประกอบด้วยส่วนสำคัญ 2 ส่วนคือ

2.1 แนวทางต่าง ๆ ที่จะใช้พิจารณาชิ้นงาน (criteria) หรืออาจเรียกว่าองค์ประกอบ หรือรายการที่จะประเมิน เช่น ในผลงานการเรียนชิ้นหนึ่ง เราอาจมีแนวทางประเมินโดยดูสิ่งต่อไปนี้ ได้แก่ จุดมุ่งหมาย (purpose) การจัดองค์ประกอบ (organization) รายละเอียด (details) ท่วงทำนอง การเขียน (voice style) หลักไวยากรณ์ตัวสะกด (mechanics) เป็นต้น

2.2 คำอธิบายถึงระดับคุณภาพของแต่ละองค์ประกอบ จึงเรียงลำดับตั้งแต่สูงสุดถึงต่ำสุด ซึ่งจะกำหนดคุณภาพเป็นกี่ยกระดับขึ้นอยู่กับระดับความเหมาะสม

นอกจากนี้ Judith (2000) กล่าวว่า เกณฑ์การให้คะแนนผลงาน (scoring rubric) มี 2 วิธีการ คือเกณฑ์การให้คะแนนเป็นภาพรวม หรือเกณฑ์รวม (holistic scoring rubric) และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ หรือเกณฑ์ย่อย (analytic scoring rubric) ดังนี้ วิธีแรก เกณฑ์การให้คะแนนเป็นภาพรวม ในเกณฑ์การให้คะแนนนี้จะแบ่งเป็นระดับคะแนนเดียว หรือระดับคุณภาพเดียวของผลงาน หรือผลการปฏิบัติโดยรวมทั้งหมด วิธีนี้จะใช้เมื่อต้องการมองเป็นคะแนนรวมมากกว่าที่จะพิจารณาจากการอธิบายคุณภาพของผลงานนั้น ต้องการดูภาพรวมของผลสัมฤทธิ์อย่างรวดเร็ว ถึงแม้จะมีมิติเดียวแต่ก็สามารถบอกถึงคุณภาพผลงานนั้นได้ ข้อเสียของเกณฑ์การให้คะแนนผลงานนี้ คือ ถ้านักเรียนสองคนได้คะแนนในระดับเดียวกัน แต่เหตุผลของการให้คะแนนจะแตกต่างกันอย่างมาก ไม่สามารถระบุจุดเด่น จุดด้อยของนักเรียนได้อย่างชัดเจน ส่วนวิธีที่สอง เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ หรือเกณฑ์ย่อยจะมีระดับคะแนนหรือระดับคุณภาพเป็นเกณฑ์ย่อยๆของผลงาน หรือผลการปฏิบัติ แต่ละคะแนนจะแสดงถึงมิติต่างๆที่สำคัญ หรือลักษณะเฉพาะของผลงานนั้น ซึ่งจะวิธีนี้จะใช้เมื่อต้องการใช้เพื่อพิจารณาคุณภาพผลงาน หรือผลการปฏิบัติงานที่ต้องการรายละเอียดชัดเจน เพื่อเป็นข้อมูลย้อนกลับแก่นักเรียน และผู้ปกครอง เป็นทักษะที่ซับซ้อน มีหลายมิติจึงต้องการให้มีความชัดเจน กระจ่างในผลงานนั้น ข้อเสียของวิธีการนี้คือ คะแนนที่ได้ช้า และใช้เวลานาน

ส่วน Ronis (2000) กล่าวว่า ในการให้คะแนนของนักเรียนมี 2 แนวทางคือ วิธีภาพรวมจะเป็นรูปแบบการประเมินที่พิจารณาตัวอย่างผลงานของนักเรียนแบบโดยรวม รูปแบบการประเมิน จะใช้ตัวเลขเดียว เช่น จาก 0 ถึง 4 หรือ 0 ถึง 6 วิธีนี้จะใช้เมื่อต้องการความรวดเร็ว

ในวิธีการให้คะแนน และอีกแนวทางหนึ่งคือ วิธีแยกองค์ประกอบ จะแยกคะแนนออกเป็นส่วนๆ โดยจะแตกต่างกันไปตามคุณภาพ หรือคุณลักษณะของผลงาน รูปแบบการให้คะแนนแบบนี้จะใช้เวลามากกว่าวิธีแบบภาพรวม แต่จะให้ข้อมูลที่มียละเอียด เพราะวิธีนี้สามารถชี้ถึงสิ่งที่จำเป็นที่ ต้องมีการปรับปรุงอย่างแม่นยำ มักจะใช้บ่อยสำหรับวินิจฉัยวัตถุประสงค์ หรือหลักสูตรและ โปรแกรมการประเมินการสอน

Kubiszyn and Borich (2000) กล่าวว่า ในการประเมินผลการปฏิบัติงานจะต้อง มีระบบการให้คะแนนที่เรียกว่า scoring rubric เป็นโมเดลคำตอบที่แสดงถึงการปฏิบัติงานของ ผู้เรียนที่ต้องการมีการเปรียบเทียบ โดยจะต้องมีการให้รายละเอียดของรายการคำตอบ หรือตัวอย่าง ของคำตอบที่เป็นที่ยอมรับ ระบบการให้คะแนนในการประเมินผลการปฏิบัติโดยทั่วไปจะประกอบด้วย รฐบรีค 3 กลุ่มได้แก่ แบบตรวจสอบรายการ (checklists) มาตรฐานประมาณค่า (rating scales) และการให้คะแนนผลรวม (holistic scoring) ซึ่งแต่ละวิธีจะมีข้อดีและข้อจำกัดใน การนำไปใช้ให้คะแนน จึงต้องเลือกให้เหมาะสมกับการปฏิบัติงานนั้น

Moskal (2000) กล่าวว่า เกณฑ์การให้คะแนน (scoring rubric) หมายถึง การบรรยาย ถึงแบบแผนการให้คะแนนที่พัฒนาขึ้นโดยครู หรือนักประเมิน เพื่อใช้เป็นแนวทางในการวิเคราะห์ ผลงาน หรือกระบวนการจากความอูตสาหะของนักเรียน (Brookhart, 1999) เกณฑ์การให้คะแนน จะนำมาใช้เมื่อต้องการตัดสินผลงานว่ามีคุณภาพอยู่ในระดับใด และอาจจะนำมาใช้ประเมิน วิชาเรียนและกิจกรรมต่างๆ โดยปกติการใช้เกณฑ์การให้คะแนนจะใช้ในการประเมินทักษะ การเขียนจะพิจารณาจากคุณภาพของตัวอย่างงานเขียน ซึ่งอาจจะมีจำนวนมากขึ้นอยู่กับเกณฑ์ (criteria) ที่ผู้ประเมินแต่ละคนเป็นผู้กำหนดขึ้นมา เกณฑ์การให้คะแนนเป็นหนึ่งในทางเลือกที่ใช้ ในการประเมินผลงานของนักเรียน โดยจะตั้งอยู่บนพื้นฐานที่ใช้ในการอธิบายแต่ละระดับและ จะประเมินตามเกณฑ์ คำอธิบายนี้จะช่วยให้นักเรียนเข้าใจหลักเกณฑ์ของคะแนน และทำให้ทราบ ว่าอะไรที่พวกเขาจำเป็นต้องพัฒนาปรับปรุงการปฏิบัติของพวกเขาในครั้งต่อไป ประโยชน์ของการ นำเกณฑ์การให้คะแนนมาใช้ในการประเมิน มีอย่างน้อย 2 ประการ กล่าวคือ 1) เป็นการสนับสนุน การสอนให้บรรลุวัตถุประสงค์ 2) เป็นการให้ข้อมูลย้อนกลับแก่นักเรียนเกี่ยวกับวิธีการปรับปรุงผล การปฏิบัติงานของพวกเขา เกณฑ์การให้คะแนน (scoring rubric) มีหลายรูปแบบ แตกต่างกันไป ขึ้น อยู่กับวัตถุประสงค์การประเมิน ในส่วนนี้เขาได้อธิบายความแตกต่างระหว่างเกณฑ์การให้คะแนน แบบภาพรวม และแยกองค์ประกอบ ดังนี้ เกณฑ์การให้คะแนนเป็นภาพรวม หรือเกณฑ์รวม (holistic scoring rubric) จะเป็นการพิจารณารวมในระดับคะแนนนั้นโดยตัดสินจากคุณภาพของ กระบวนการหรือผลงานนั้น และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ หรือเกณฑ์ย่อย (analytic scoring rubric) จะแยกการประเมินในแต่ละองค์ประกอบ ซึ่งแต่ละองค์ประกอบ จะมี คะแนนที่อธิบายในระดับที่แตกต่างกัน

Herman et al. (1996) กล่าวว่า วิธีการแบบภาพรวม เป็นวิธีการที่ครูหรือผู้ประเมินจะให้คะแนนเดียว (single score) แสดงถึงการปฏิบัติงานทั้งหมดของนักเรียน ส่วนวิธีการแบบแยกองค์ประกอบ จะแยกออกเป็นแต่ละระดับที่แตกต่างกันของการปฏิบัติงาน ระบุปริมาณมากและวิธีการให้คะแนนจะถูกนำมาผสมกันในการตัดสินแบบภาพรวม และแบบแยกองค์ประกอบ วิธีการแบบภาพรวมจะใช้เวลาน้อยกว่าแบบแยกองค์ประกอบ ถ้าต้องการข้อมูลย้อนกลับในการพัฒนาการเรียนรู้อันวิธีการแบบแยกองค์ประกอบ จะนำมาใช้และมีประโยชน์มากกว่าอีกวิธีหนึ่งซึ่งวิธีการนี้จะให้ผลโดยตรงที่ชัดเจนแก่ผู้เรียน ผู้ปกครอง และครู ในการตัดสินใจการปฏิบัติงาน

Nitko (2000) กล่าวถึง เกณฑ์การให้คะแนน ว่าเป็นกฎเกณฑ์หรือหลักที่ใช้ในการประเมินคุณภาพการปฏิบัติงานของนักเรียน โดยกฎเกณฑ์นี้อาจอยู่ในรูปของมาตรฐานค่า หรือแบบตรวจสอบรายการ มาตรฐานจะประกอบด้วย ตัวเลขหลายตัวที่ใช้บ่อยคือ 0 ถึง 3 หรือ 1 ถึง 4 จะเป็นการบอกถึงระดับคุณภาพของการปฏิบัติ ตัวเลขที่ใช้จะแสดงถึงลำดับของระดับคุณภาพ เช่น ตัวเลข 4 หมายถึง ระดับสูงสุด (ดีมาก) ตัวเลข 3 หมายถึง ระดับรองลงมา (ดี) ลดหลั่นจนถึงระดับต่ำสุด จากที่กล่าวมาจะพบว่า การให้คะแนนการปฏิบัติของนักเรียนแต่ละคนคุณภาพของการปฏิบัติที่ระดับต่างๆจะบรรยายด้วยเกณฑ์การให้คะแนน นอกจากนี้เขาได้แบ่งเกณฑ์การให้คะแนน เป็น 2 แบบ คือ เกณฑ์ทั่วไป (general scoring rubric) และเกณฑ์เฉพาะเจาะจง (specific scoring rubric) โดยมีรายละเอียดดังนี้

1. เกณฑ์ทั่วไป (general scoring rubric) เป็นเกณฑ์การให้คะแนนที่ไม่ได้มีการมุ่งเน้นในงานนั้น โดยจะเป็นกรอบทั่วไปในการพัฒนาเกณฑ์ที่เฉพาะเจาะจง

2. เกณฑ์เฉพาะเจาะจง (specific scoring rubric) เป็นเกณฑ์การให้คะแนนที่พัฒนาจากเกณฑ์ทั่วไปในการให้คะแนน แต่ละส่วนของงานนั้น ในการใช้เกณฑ์เฉพาะเจาะจงนี้ในการประเมินนักเรียน ต้องแน่ใจว่าเกณฑ์นี้เป็นส่วนหนึ่งในกรอบการให้คะแนนทั่วไป

นอกจากนี้เขายังได้แบ่งประเภทของเกณฑ์การให้คะแนน เป็น 3 ประเภทดังนี้

1. holistic scoring rubric เกณฑ์ประเภทนี้จะเป็นมาตรวัด (scale) หรือคะแนนที่แสดงผลงานหรือกระบวนการโดยรวมทั้งหมด ซึ่งจะไม่แยกคะแนนออกเป็นแต่ละองค์ประกอบหรือรายการ วิธีการนี้ใช้ง่ายและประหยัดเวลา

2. analytic scoring rubric เกณฑ์นี้จะใช้มาตรวัดหรือคะแนนที่แยกลักษณะของผลงานหรือกระบวนการออกเป็นแต่ละองค์ประกอบหรือรายการโดยคะแนนที่ได้จากแต่ละองค์ประกอบหรือรายการ จะต้องรวมคะแนนเป็นคะแนนรวมทั้งหมดของผลงาน วิธีนี้จะใช้เวลามาก ในรายละเอียดของแต่ละระดับจะมีประโยชน์เมื่อต้องการวินิจฉัยหรือช่วยนักเรียนให้เข้าใจถึงสิ่งที่คาดหวังจากการประเมินที่ได้จากการประเมิน จะใช้เวลามากแต่ก็มีประโยชน์อย่างยิ่งที่จะช่วยในการเรียนรู้ของนักเรียน

3. annotated holistic scoring Rubric ผู้ประเมินจะใช้มาตรฐานวัดรวม ในการประเมินก่อน ซึ่งมาตรฐาน หรือคำอธิบายจะแสดงลักษณะของจุดเด่น (strengths) และจุดด้อย (weakness) ที่สนับสนุน มาตรฐานวัดรวม วิธีการนี้เกิดจากการรวมเกณฑ์แบบภาพรวม และแบบแยกองค์ประกอบ ไว้ด้วยกัน เมื่อประเมินการปฏิบัติโดยภาพรวมจะต้องเลือกหนึ่งหรือสองประเด็นในการอธิบายหรือ ใช้มาตรฐานวัดแบบแยกองค์ประกอบ ข้อดีของวิธีการนี้จะทำให้มีความรวดเร็วในการให้คะแนนและ เป็นการสนับสนุนว่าในแต่ละระดับนั้นจะให้ข้อมูลย้อนกลับแก่นักเรียนได้

จากที่กล่าวมาข้างต้น ได้มีผู้ให้คำจำกัดความแตกต่างกันไปสรุปได้ว่า เกณฑ์การให้คะแนน (scoring rubric) มีลักษณะเป็นระดับผลการทำงาน ซึ่งในแต่ละระดับจะมีรายการที่บรรยายถึงประเด็นที่ทำการประเมิน เพื่อให้ทราบว่าผู้ถูกประเมินทำงานอย่างไร ผลงานเป็นอย่างไรและอยู่ในระดับใด โดยทั่วไปจะเสนอ เป็น 2 รูปแบบ คือ 1) เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) เป็นเกณฑ์การให้คะแนนที่พิจารณาโดยรวมไม่มีการแยกประเด็นย่อยๆ การให้คะแนนขึ้นอยู่กับคุณภาพของผลงาน 2) เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) เป็นเกณฑ์การให้คะแนนที่แยกเป็นประเด็นย่อยๆ ตามองค์ประกอบและกำหนดน้ำหนักคะแนนในแต่ละส่วน

ตอนที่ 3 ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

Cronbach และ คณะ (1972) ได้เสนอทฤษฎีสำหรับการวิเคราะห์ความน่าเชื่อถือของผลการวัดพฤติกรรมในสถานการณ์ต่างๆ ซึ่งต่อมารู้จักกันในชื่อของ ทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

ทฤษฎีการสรุปอ้างอิง (Generalizability Theory) หรือ G-Theory เป็นทฤษฎีทางสถิติของการวิเคราะห์ความน่าเชื่อถือของผลการวัด ในสถานการณ์ของการวัดผลลักษณะต่างๆที่เป็นเป้าหมายของการนำเครื่องมือไปใช้ ตามแนวความคิดของ G-Theory คะแนนที่ได้จากการวัดภายใต้สถานการณ์ หรือเงื่อนไขเดียวด้วยการสอบชุดเดียว และทำการสอบครั้งเดียว คะแนนที่ได้จะไม่สามารถเชื่อถือได้อย่างเต็มที่ เพราะคะแนนที่ได้จะไม่เป็นตัวแทนที่ดีของคะแนนจริง ซึ่งเป็นคะแนนเฉลี่ยที่ได้จากการทดสอบในหลายๆสถานการณ์ด้วยแบบสอบหลายๆชุด และทำการสอบหลายๆครั้ง แต่ G-Theory ได้พิจารณาถึงแหล่งความคลาดเคลื่อนพหุ (multiple sources of error) ที่สามารถวิเคราะห์ความคลาดเคลื่อนจากหลายแหล่งพร้อมกันไป ผู้บริหารการทดสอบจึงสามารถตัดสินใจได้ว่าควรใช้สถานการณ์หรือเงื่อนไขการวัดแบบใดจึงจะได้คะแนนที่เชื่อถือถึงระดับที่ต้องการ ผลการวิเคราะห์ด้วย G-Theory จะให้ค่าสัมประสิทธิ์เชิงสรุปที่แสดงถึงระดับความน่าเชื่อถือของคะแนนที่ได้จากการวัด (level of dependability) เรียกว่า สัมประสิทธิ์การสรุปอ้างอิงความน่าเชื่อถือของผลการวัด (Generalizability Coefficient) ซึ่งคล้ายกับสัมประสิทธิ์ความเที่ยง (Reliability Coefficient) ในทฤษฎีการทดสอบแบบดั้งเดิม (ศิริชัย กาญจนวาสี, 2542)

หลังจากนั้น ได้มีนักการศึกษาหลายคนได้ศึกษาทฤษฎีการสรุปอ้างอิงนี้ รวมทั้งได้ปรับปรุงวิธีการให้รัดกุมและง่ายขึ้น ผู้ที่มีส่วนในการพัฒนาทฤษฎีนี้อย่างมากคือ Brennan (1983) ได้ปรับปรุงวิธีการให้ง่ายขึ้นในแง่การตีความและการคำนวณ ต่อมา Cardinet et al. (1976, 1981, 1983) ได้ขยายความในทฤษฎีบางจุดเพื่อให้สามารถประยุกต์ใช้ได้กว้างขวางกว่าเดิม Cardinet and Allal (1983) กล่าวถึงการประยุกต์ทฤษฎีการสรุปอ้างอิงว่า การประยุกต์ใช้ในทางจิตวิทยานั้น นิยมใช้กับการศึกษาถึงความแตกต่างระหว่างบุคคล สิ่งที่ถูกวัดส่วนใหญ่ ได้แก่ บุคคล องค์ประกอบของการวัดจึงจำกัดอยู่เฉพาะแต่องค์ประกอบที่สอดคล้องกับเงื่อนไขการวัดเท่านั้น ตามหลักการแล้วทฤษฎีนี้สามารถทำได้กว้างขวางยิ่งขึ้น ดังนั้น Cardinet et al. (1981) จึงนำหลักการของความสมมาตร (Principle of Symmetry) มาใช้ในการวิเคราะห์การสรุปอ้างอิง ซึ่งหลักการนี้ได้รับการยอมรับว่าองค์ประกอบแต่ละตัวของรูปแบบการวัดใดๆ สามารถเลือกให้เป็นสิ่งที่ถูกวัดได้ทั้งสิ้น ทั้งนี้ขึ้นอยู่กับจุดมุ่งหมายของการวัดนั้นๆ ต่อมา Vann der Kamp (1976) ได้อธิบายเพิ่มเติมเกี่ยวกับทฤษฎีการสรุปอ้างอิงนี้ว่า สามารถใช้อธิบายค่าความเที่ยงได้ทุกสถานการณ์ โดยเฉพาะอย่างยิ่งการสอบที่ไม่ใช่ข้อสอบแบบเลือกตอบ เนื่องจากทฤษฎีนี้ไม่ได้กำหนดข้อตกลงเกี่ยวกับคุณสมบัติของความเท่าเทียม หรือความเป็น

คูชานานของการวัด ผู้ตรวจข้อสอบแต่ละคนก็ไม่จำเป็นต้องมีคุณสมบัติของความเท่าเทียม ทฤษฎีนี้ถือว่าคะแนนจริงคือ คะแนนเอกภพ (Universe Score) ซึ่งเป็นค่าเฉลี่ยของคะแนนสังเกต ที่เป็นไปได้ทั้งหมดภายในเอกภพที่ผู้วัดสนใจ คะแนนเอกภพของผู้สอบแต่ละคนมิได้หลายค่า เช่นเดียวกับค่าสัมประสิทธิ์การสุรูปอ้างอิง ทั้งนี้ขึ้นอยู่กับเอกภพที่ผู้วัดหรือผู้ประเมินสนใจ (กนกวรรณ เตียมชัย, 2539)

ความรู้พื้นฐานเกี่ยวกับโครงสร้างของการศึกษาทฤษฎีการสุรูปอ้างอิง (Brennan, 1983) ได้เสนอคำศัพท์ที่เฉพาะในทฤษฎีการสุรูปอ้างอิง ไว้ดังนี้

1. องค์ประกอบ (Facet) หมายถึง ชุดของเงื่อนไขการวัดที่คล้ายคลึงกัน (Similar Conditions of Measurement) เช่น องค์ประกอบของข้อสอบ (Item Facet) องค์ประกอบของผู้ตรวจข้อสอบ (Rater Facet) เป็นต้น

2. เงื่อนไขการวัด (Condition) หมายถึง ระดับขององค์ประกอบที่ทำให้ได้ค่าสังเกตแต่ละค่าขึ้นมาในการวัดครั้งหนึ่ง ๆ อาจกล่าวได้ว่า เป็นขนาดของกลุ่มตัวอย่างขององค์ประกอบนั้น เช่น ข้อสอบแต่ละข้อเป็นเงื่อนไขการวัดหนึ่ง ๆ ขององค์ประกอบแบบทดสอบ

3. รูปแบบความสัมพันธ์ระหว่างองค์ประกอบ หมายถึง องค์ประกอบที่เป็นสิ่งที่ชี้ถึงแหล่งความแปรปรวนของรูปแบบการวัด ในแบบจำลองการวิเคราะห์ความแปรปรวน ซึ่งรูปแบบความสัมพันธ์นี้ แบ่งออกเป็น 3 ลักษณะ คือ

3.1 ความสัมพันธ์แบบไขว้ (Crossed) หมายถึง ความสัมพันธ์แต่ละลักษณะที่แต่ละระดับของสิ่งที่ถูกวัด ถูกวัดภายใต้เงื่อนไขเดียวกันทั้งหมด สัญลักษณ์คือ "x" อ่านว่า "crossed with" เช่น กำหนดรูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็น $p \times i \times r$ หมายถึง นักเรียน (p) ทำข้อสอบ (i) ทุกข้อ และผู้ตรวจ (r) ตรวจข้อสอบของนักเรียนทุกคน ทุก ๆ ข้อ

3.2 ความสัมพันธ์แบบแฝง (Nested) หมายถึง ความสัมพันธ์ในลักษณะที่แต่ละระดับของสิ่งที่ถูกวัด ถูกวัดภายใต้เงื่อนไขที่แตกต่างกัน สัญลักษณ์คือ ":" อ่านว่า "nested within" เช่น รูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็น $i:r$ หมายถึง ผู้ตรวจข้อสอบ (r) ตรวจข้อสอบ (i) ต่างข้อกัน

3.3 ความสัมพันธ์แบบผสม (Confounded) หมายถึง ความสัมพันธ์ที่มีทั้งความสัมพันธ์แบบไขว้และความสัมพันธ์แบบแฝงปนกันอยู่ เช่น $p \times (i:t)$ หมายถึง นักเรียน (p) ทำข้อสอบ (i) บางข้อที่อยู่ในแบบทดสอบ (t)

4. เอกภพ (Universe) หมายถึง เงื่อนไขของการวัดทั้งหมดของแต่ละองค์ประกอบ เช่น จำนวนข้อสอบทุกข้อ จำนวนผู้ตรวจทั้งหมด นักเรียนทุกคน เอกภพมีความหมายคล้ายกับ คำว่า ประชากร (Population) คำว่าเอกภพนิยมใช้กับองค์ประกอบที่เกี่ยวกับการวัด ส่วนคำว่า ประชากรใช้กับองค์ประกอบที่เป็นสิ่งที่ถูกวัด

คะแนนเอกภพ (Universe Scores) หมายถึง คะแนนของเงื่อนไขการวัดทั้งหมดของแต่ละองค์ประกอบซึ่งคล้ายกับคะแนนจริง (Ideal Score) ของสิ่งที่วัดในทฤษฎีการวัดแบบดั้งเดิม

5. เอกภพของการสังเกตที่ยอมรับได้ (Universe of Admissible Observation) หมายถึง กลุ่มเงื่อนไขของการวัดที่สามารถวัดหรือสังเกตได้ในแต่ละองค์ประกอบ เช่นต้องการวัดผล การเรียนของนักเรียน ด้วยการที่ครูต้องการข้อสอบที่สามารถนำไปใช้ได้ และต้องการผู้ตรวจที่มี คุณภาพ ดังนั้น ครูจึงกำหนดเงื่อนไขของการวัดเป็นองค์ประกอบของข้อสอบ และองค์ประกอบ ของผู้ตรวจให้คะแนน ซึ่งเรียกว่าเอกภพของการสังเกตที่ยอมรับได้ของครูคนนี้ ประกอบด้วย องค์ประกอบของข้อสอบ และองค์ประกอบของผู้ตรวจให้คะแนน

6. เอกภพของการสรุปอ้างอิง (Universe of Generalization) หมายถึงเงื่อนไขการวัด ทั้งหมดขององค์ประกอบที่ผู้วิจัยต้องการสรุปอ้างอิงผลการวัด จากกลุ่มตัวอย่างของเงื่อนไขการวัด เหล่านี้ไปยัง กลุ่มเงื่อนไขการวัดทั้งหมดขององค์ประกอบนั้นๆ ดังนั้น เงื่อนไขขององค์ประกอบ ในเอกภพของการสรุปอ้างอิง จึงเป็นส่วนย่อยๆ (Subset) ในเอกภพของการสังเกตที่ยอมรับได้ เอกภพของการสรุปอ้างอิงนี้ ผู้วิจัยต้องระบุประเด็นต่างๆ ให้ชัดเจน ผลที่ได้นำไปใช้ในการ ออกแบบการวัด ในการวิเคราะห์เพื่อการตัดสินใจ (D Study) ต่อไป

7. การศึกษาเพื่อการสรุปอ้างอิง (Generalizability Study หรือ G Study) หมายถึง การวิเคราะห์การสรุปอ้างอิง เพื่อประมาณค่าความแปรปรวนจากแหล่งต่างๆ ของรูปแบบการวัด เริ่มจากการกำหนดองค์ประกอบต่างๆที่ต้องการวัด กำหนดความสัมพันธ์ขององค์ประกอบและ กำหนดเอกภพของการสังเกตที่ยอมรับได้ และใช้วิธีการวิเคราะห์ความแปรปรวน การประมาณค่า ความแปรปรวนจากแหล่งต่างๆ

8. การศึกษาเพื่อการตัดสินใจ (Decision Study หรือ D Study) หมายถึง การศึกษาเพื่อ ประเมินวิธีการวัดที่เหมาะสม เพื่อการตัดสินใจหรือลงข้อสรุป เป็นขั้นสุดท้ายของการวิเคราะห์ การสรุปอ้างอิง เริ่มจากผู้วิจัยต้องระบุสิ่งที่ต้องการวัด เอกภพของการสรุปอ้างอิง จำนวนเงื่อนไข ของการวัด หรือขนาดของกลุ่มตัวอย่างของแต่ละองค์ประกอบ และรูปแบบของการวัด แล้วหา ค่าความแปรปรวนจากการศึกษา ประมาณค่าความแปรปรวน หาค่าสัมประสิทธิ์การสรุปอ้างอิง หรือความเที่ยงของแบบทดสอบ และขั้นสุดท้ายคือปรับรูปแบบการวัด เพื่อหาวิธีที่เหมาะสม เพื่อให้เกิดประโยชน์สูงสุดต่อไป

ขนาดของกลุ่มตัวอย่างในการศึกษาเพื่อการตัดสินใจ ไม่จำเป็นต้องมีขนาดเหมือนกับ ขนาดของกลุ่มตัวอย่างที่ใช้ในการศึกษาเพื่อการสรุปอ้างอิง ทั้งนี้ขึ้นอยู่กับนำไปใช้เพื่อ การตัดสินใจเกี่ยวกับสิ่งที่ต้องการวัดจริงๆ ซึ่งถูกกำหนดขึ้นโดยมีขนาดของกลุ่มตัวอย่างใน การศึกษาเพื่อการสรุปอ้างอิงเป็นพื้นฐาน

9. ความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์ (Absolute Error Variance หรือ $\sigma^2(\Delta)$) หมายถึง ความแปรปรวนของผลต่างระหว่างคะแนนสังเกตกับคะแนนเอกภพ ซึ่งมีค่าเท่ากับผลรวมของความแปรปรวนที่ประมาณได้ทั้งหมด ยกเว้นความแปรปรวนของแหล่งบุคคล

10. ความแปรปรวนของความคลาดเคลื่อนแบบสัมพัทธ์ (Relative Error Variance หรือ $\sigma^2(\delta)$) หมายถึง ความแปรปรวนของผลต่างระหว่างส่วนเบี่ยงเบนของคะแนนสังเกต (ความแตกต่างของคะแนนสังเกต จากคะแนนเฉลี่ยของประชากรของคะแนนสังเกต) และส่วนเบี่ยงเบนของคะแนนเอกภพ (ความแตกต่างของคะแนนเอกภพจากคะแนนเฉลี่ยของประชากรของคะแนนเอกภพ)

11. สัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient หรือ $E\rho^2$) หมายถึง อัตราส่วนของความแปรปรวนของคะแนนเอกภพ และความแปรปรวนของคะแนนสังเกตที่คาดหวังซึ่งค่าสัมประสิทธิ์การสรุปอ้างอิงที่ประมาณค่าได้จากกำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนเอกภพและคะแนนสังเกต

11.1 สัมประสิทธิ์การสรุปอ้างอิง สำหรับการวัดที่มีการตัดสินใจในเชิงเปรียบเทียบ หรือการวัดแบบอิงกลุ่ม ความคลาดเคลื่อนในการวัดเป็นความคลาดเคลื่อนแบบสัมพัทธ์ ดังนั้น สัมประสิทธิ์การสรุปอ้างอิงประมาณค่าได้จากอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพ และผลบวกของความแปรปรวนของคะแนนเอกภพกับความแปรปรวนของความคลาดเคลื่อนสัมพัทธ์

11.2 สัมประสิทธิ์การสรุปอ้างอิง สำหรับการวัดที่มีการตัดสินใจแบบสัมบูรณ์หรือการวัดแบบอิงเกณฑ์ ความคลาดเคลื่อนในการวัดเป็นความคลาดเคลื่อนแบบสัมบูรณ์ ค่าสัมประสิทธิ์การสรุปอ้างอิง ประมาณค่าได้จาก อัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพ และผลบวกของความแปรปรวนของคะแนนเอกภพกับความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์

หลักพื้นฐานของทฤษฎีการสรุปอ้างอิง

ข้อตกลงเบื้องต้นของทฤษฎีการสรุปอ้างอิง อาศัยข้อตกลงของแบบการวิเคราะห์ความแปรปรวน ที่สัมพันธ์กับรูปแบบของการวัดที่ต้องการศึกษา โดยทั่วไปมีข้อตกลงดังนี้ (ไพรัตน์ วงษ์นาม, 2533)

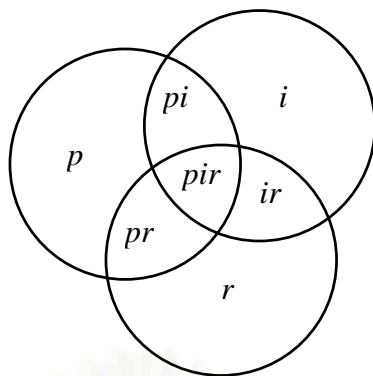
1. ต้องระบุเอกภพที่ต้องการสรุปอ้างอิงให้ชัดเจน จนสามารถบอกได้ว่ามีเงื่อนไขใดบ้างที่เป็นสมาชิกของเอกภพนั้น

2. เงื่อนไขการวัดเป็นอิสระต่อกัน กล่าวคือ คะแนนของนักศึกษาที่ทำข้อสอบข้อ i ถูกหรือผิด ไม่ขึ้นอยู่กับคำตอบข้ออื่น

3. คะแนนสังเกต (X) สมมติให้เป็นค่าการวัดในมาตราช่วง (Interval Scale)

แนวคิดในการประมาณค่าพารามิเตอร์ ของทฤษฎีการสุ่มอย่างอิง Cronbach et al. (1972) และ Brennan (1983) เสนอวิธีการประมาณค่าพารามิเตอร์ เป็น 2 ขั้นตอน คือ ขั้นที่ 1 การศึกษาเพื่อการสุ่มอย่างอิง (Generalizability Study หรือ G Study) และขั้นที่ 2 การศึกษาเพื่อการตัดสินใจ (Decision Study หรือ D Study) แต่ละขั้นมีรายละเอียดดังนี้ (ไพรัตน์ วงษ์นาม, 2533)

ขั้นที่ 1 การศึกษาเพื่อการสุ่มอย่างอิง (G Study) เป็นการประมาณค่าความแปรปรวนจากแหล่งต่างๆ ภายใต้เงื่อนไขการวัดที่ยอมรับได้ เริ่มด้วยการกำหนดสิ่งที่ถูกวัดเป็นคน (นักเรียน) หรือกลุ่มคน เช่น นักเรียนทั้งห้อง หรือทั้งโรงเรียน ให้องค์ประกอบอื่นๆ เป็นองค์ประกอบภายใต้การสังเกต เช่น ข้อสอบ ผู้สอบ ผู้ตรวจ ฯลฯ กำหนดความสัมพันธ์ระหว่างองค์ประกอบเป็นแบบไขว้ (Crossed) หรือแบบแฝง (Nested) แล้วใช้วิธีการวิเคราะห์ความแปรปรวน (Analysis of Variance, ANOVA) ที่สอดคล้องกับแบบแผนและความสัมพันธ์ขององค์ประกอบประมาณค่าเฉลี่ยกำลังสอง (Mean Square) แล้วจึงประมาณค่าความแปรปรวนของแหล่งต่างๆ เรียกค่าประมาณความแปรปรวนของสิ่งที่ถูกวัดว่า ค่าประมาณความแปรปรวนของคะแนนเอกภพ ซึ่งสอดคล้องกับ ค่าประมาณความแปรปรวนของคะแนนจริง (True Score Variance) ของทฤษฎีการวัดแบบดั้งเดิมเรียกค่าประมาณความแปรปรวนขององค์ประกอบอื่นๆ ว่า ค่าประมาณความแปรปรวนของความคลาดเคลื่อน (Estimated Error Score Variance) การวิเคราะห์ในขั้นการศึกษาเพื่อการสุ่มอย่างอิง ใช้ค่าขนาดของกลุ่มตัวอย่างที่ศึกษาจริงๆ เช่น ถ้าครูต้องการหาวิธีประเมินผลความสามารถในการเขียนเรียงความของนักเรียน ต้องเริ่มต้นจากการระบุลักษณะของข้อสอบที่ใช้วัด กำหนดผู้ตรวจข้อสอบที่มีความเชี่ยวชาญ เช่น ถ้ากำหนดว่าผู้ตรวจ (r) แต่ละคนต้องตรวจข้อสอบ (i) ทุกข้อเหมือนกัน เอกภพของการสังเกตที่ยอมรับได้ จึงเป็นแบบไขว้ใช้สัญลักษณ์ว่า $i \times r$ อ่านว่า i crossed with r จากนั้นครูต้องเก็บรวบรวมและวิเคราะห์ข้อมูลโดยสุ่มนักเรียนมา n_p คน ใช้ข้อสอบจำนวน n_i ข้อ เมื่อสอบเสร็จก็สุ่มผู้ตรวจมาตรวจข้อสอบให้คะแนนของนักเรียน โดยที่ผู้ตรวจทุกคนตรวจข้อสอบของนักเรียนทุกคนและตรวจทุกข้อแล้วนำข้อมูลมาคำนวณค่าความแปรปรวนขององค์ประกอบต่างๆ วิธีการนี้ก็คือ การศึกษาในขั้นตอน G Study ใช้สัญลักษณ์ว่า $p \times i \times r$ แบบแผนการวิเคราะห์นี้ จึงมีแหล่งความแปรปรวนทั้งหมด 7 แหล่ง Cardinet et al. (1972) ได้ตั้งชื่อแบบแผนนี้ว่า แบบแผน 7 (VII Design) เขียนแทนด้วยแผนภูมิ Venn ได้ดังภาพที่ 1



แผนภาพ 1 แผนภูมิแสดงองค์ประกอบความแปรปรวนของรูปแบบ $p \times i \times r$ เมื่อตัวประกอบที่ศึกษาทั้งหมดเป็นตัวประกอบสุ่ม

จากแผนภาพ 1 จะเห็นว่าผลหลัก (Main Effects) อยู่ 3 ค่าคือ 1) ผลของผู้สอบ (p) 2) ผลของข้อสอบ (i) และ 3) ผลของผู้ตรวจ (r) ส่วนผลร่วมหรือปฏิสัมพันธ์สองระดับมี 3 ค่าคือ 1) ผลร่วมของผู้สอบและข้อสอบ (pi) 2) ผลร่วมของผู้สอบและผู้ตรวจ (pr) 3) ผลร่วมของข้อสอบและผู้ตรวจ (ir) และผลร่วมทั้งหมดของผู้สอบ ข้อสอบและผู้ตรวจ (pir)

ค่าความแปรปรวนของแหล่งต่างๆเหล่านี้ คำนวณมาจากค่าเฉลี่ยกำลังสอง (Mean Square) โดยใช้การวิเคราะห์ความแปรปรวนแบบ Factorial Design $p \times i \times r$ ในการประมาณค่าความแปรปรวนจากค่าเฉลี่ยกำลังสองนั้น ใช้สูตรเฉพาะอย่างที่สอดคล้องกับรูปแบบ

สูตรคำนวณค่าเฉลี่ยของกำลังสองของความเบี่ยงเบน (Means Square of Deviation หรือ $MS(\alpha)$) 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

และจากความสัมพันธ์ $p \times i \times r$ หาค่าความแปรปรวนได้จากสูตรต่อไปนี้ (Brennan, 1983)

$$\begin{aligned}\sigma^2(pir) &= MS(pir) \\ \sigma^2(pi) &= 1/n_r [MS(pi) - MS(pir)] \\ \sigma^2(pr) &= 1/n_i [MS(pr) - MS(pir)] \\ \sigma^2(ir) &= 1/n_p [MS(ir) - MS(pir)] \\ \sigma^2(r) &= 1/n_p n_i [MS(r) - MS(pr) - MS(ir) + MS(pir)] \\ \sigma^2(i) &= 1/n_p n_r [MS(i) - MS(ir) - MS(pi) + MS(pir)] \\ \sigma^2(p) &= 1/n_i n_r [MS(p) - MS(pi) - MS(pr) + MS(pir)]\end{aligned}$$

ขั้นที่ 2 การศึกษาเพื่อการตัดสินใจ (D Study) ขั้นนี้เน้นการใช้ค่าประมาณ และการตีความองค์ประกอบความแปรปรวน สำหรับการตัดสินใจภายใต้วิธีการวัดที่เหมาะสม มีประเด็นสำคัญ ดังนี้

2.1 เอกภพของการสรุปอ้างอิง เป้าหมายสำคัญของการศึกษาเพื่อการตัดสินใจ ได้แก่ การกำหนดลักษณะเฉพาะของเอกภพของการสรุปอ้างอิง ที่ผู้ตัดสินใจต้องการสรุปอ้างอิง ถึงอาจประกอบด้วยเงื่อนไขทั้งหมด ในเอกภพของการสังเกตที่ยอมรับได้ หรืออาจเป็นเซตย่อย (Subset) ของเอกภพการสังเกตที่ยอมรับได้

2.2 ขนาดของตัวอย่างการศึกษาเพื่อการตัดสินใจ จำนวนเงื่อนไขขององค์ประกอบในการศึกษาเพื่อการตัดสินใจ สามารถกำหนดให้แตกต่างจาก จำนวนเงื่อนไขในการศึกษาเพื่อการสรุปอ้างอิง โดยใช้สัญลักษณ์ (') แทนขนาดตัวอย่างของการศึกษาเพื่อการตัดสินใจ เช่น n'_i แทนจำนวนข้อสอบ และ n'_r แทนจำนวนผู้ตรวจ

2.3 โครงสร้างแบบแผนการศึกษาเพื่อการตัดสินใจ นอกจากระบุขนาดของตัวอย่างในการศึกษาเพื่อการตัดสินใจ ต้องระบุรูปแบบโครงสร้างของแบบแผน หรือความสัมพันธ์ขององค์ประกอบที่ต้องการศึกษา เช่น ถ้าต้องการตัดสินใจว่าในการสอบครั้งนั้น นักเรียนต้องทำข้อสอบเหมือนกันทั้ง n_i ข้อ และผู้ตรวจข้อสอบทั้ง n_r คน ต้องตรวจข้อสอบทุกๆ ข้อ แบบแผนดังกล่าวนี้เป็น $p \times I \times R$ แต่ไม่มีความจำเป็นที่ต้องทำเช่นนั้นเสมอไป ซึ่งอาจตัดสินใจว่าผู้เข้าสอบทุกคนทำข้อสอบทุกข้อ แต่ผู้ตรวจแต่ละคนตรวจคำตอบต่างข้อกัน แบบแผนการศึกษาเพื่อการตัดสินใจเป็น $p \times (I:R)$

2.4 การประมาณค่าความแปรปรวน ในขั้นการศึกษาเพื่อการตัดสินใจ การวิเคราะห์ข้อมูลในขั้นการศึกษาเพื่อการตัดสินใจ ขึ้นอยู่กับการตัดสินใจของนักวัดผลหรือนักวิจัย ดังนั้นต้องมีการประมาณค่าความแปรปรวนขึ้นมาใหม่อีกครั้ง โดยอาศัยผลจากการประมาณค่าในขั้นการศึกษาเพื่อการสรุปอ้างอิงเป็นฐาน และให้สอดคล้องกับแบบแผนและขนาดของตัวอย่างที่ต้องการตัดสินใจ

2.5 ประมาณค่าความแปรปรวนของความคลาดเคลื่อน นอกจากการประมาณค่าองค์ประกอบความแปรปรวน (Variance Component) ตามรูปแบบและขนาดของกลุ่มตัวอย่างที่ต้องการแล้ว ต้องประมาณค่าความแปรปรวนของความคลาดเคลื่อนแบ่งเป็น 2 ชนิด คือ

2.5.1 ความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์ (Absolute Error Variance) ใช้สัญลักษณ์ $\sigma^2(\Delta)$ และค่าประมาณใช้สัญลักษณ์ $\hat{\sigma}^2(\Delta)$ เป็นความแปรปรวนของความแตกต่างระหว่างคะแนนสังเกตกับคะแนนเอกภพของผู้เข้าสอบ คำนวณจากผลบวกของค่าองค์ประกอบความแปรปรวนอื่นๆ ทั้งหมด ยกเว้นความแปรปรวนของคะแนนเอกภพ

$$\sigma^2(\Delta) = \sigma^2(I) + \sigma^2(R) + \sigma^2(pI) + \sigma^2(pR) + \sigma^2(IR) + \sigma^2(pIR)$$

2.5.2 ความแปรปรวนของความคลาดเคลื่อนแบบสัมพัทธ์ (Relative Error Variance) ใช้สัญลักษณ์ $\sigma^2(\delta)$ และค่าประมาณใช้สัญลักษณ์ $\hat{\sigma}^2(\delta)$ คำนวณจากผลบวกค่าขององค์ประกอบความแปรปรวนของผลรวมระหว่างสิ่งที่ถูกวัด (p) กับพาเซ็ทอื่น ๆ

$$\sigma^2(\delta) = \sigma^2(pI) + \sigma^2(pR) + \sigma^2(pIR)$$

2.6 ประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เป็นขั้นตอนสุดท้ายในการทำการศึกษาเพื่อการตัดสินใจ คือการคำนวณค่าสัมประสิทธิ์การสรุปอ้างอิง ซึ่งเป็นดัชนีบ่งถึงความเชื่อถือได้ของการวัด (Dependability of Measurement) ใช้สูตรคำนวณตามนิยามความเชื่อมั่นของทฤษฎีการวัดแบบดั้งเดิมในรูปของ Intraclass Correlation ส่วนค่าขนาดของความแปรปรวนของคะแนนเอกภพสิ่งที่วัด ก็คือค่าที่ใช้แทนความแปรปรวนของคะแนนจริงในทฤษฎีการทดสอบแบบดั้งเดิม ดังนี้ (Brennan and Kane, 1977 อ้างถึงใน ไพรัตน์ วงษ์นาม, 2533)

$$E\rho^2 = \frac{\sigma^2(p)}{\sigma(p) + \sigma^2(\Delta)} \quad \text{หรือ}$$

$$E\rho^2 = \frac{\sigma^2(p)}{\sigma(p) + \sigma^2(\delta)}$$

เมื่อ $E\rho^2$ คือค่าสัมประสิทธิ์การสรุปอ้างอิง มีความหมายเดียวกับค่าความเที่ยง



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตอนที่ 4 งานวิจัยอื่นๆที่เกี่ยวข้อง

งานวิจัยต่างประเทศ

Suzanne Lane et al. (1996) ได้ศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิงและความตรงของการประเมินผลงานของนักเรียนในวิชาคณิตศาสตร์ที่พัฒนาจากเครื่องมือประเมินความรู้แบบ QUASAR ชื่อว่า QCAI (QUASAR Cognitive Assessment Instrument) เป็นชุดของข้อสอบแบบปลายเปิด โดยเน้นวิธีการตรวจให้คะแนนแบบประเมินภาพรวม (Holistic) และเกณฑ์การให้คะแนนแบบรูบริก (rubric scoring) ใช้ทฤษฎีการสรุปอ้างอิงตรวจสอบความสอดคล้องภายในข้อสอบและภายในผู้ตรวจ โดยมีแบบแผน 3 รูปแบบ คือ 1) pxt ผู้สอบทุกคนทำข้อสอบทุกข้อ 2) pxtxr ผู้สอบทุกคนทำข้อสอบทุกข้อ และผู้ตรวจแต่ละคนตรวจข้อสอบทุกข้อ 3) (p:s)xt ผู้สอบที่แบ่งในโรงเรียนทำข้อสอบทุกข้อ ผลการศึกษาพบว่า ความคลาดเคลื่อนเนื่องจากผู้ตรวจมีน้อยมาก เนื่องจากการให้รายละเอียดเกี่ยวกับเกณฑ์การให้คะแนนแบบรูบริก และตัวอย่างการตอบของนักเรียน รวมทั้งการฝึกฝนวิธีการตรวจให้คะแนนกับผู้ตรวจถูกต้องเพียงพอ มีความแตกต่างที่ต้องคำนึงในการทำข้อสอบของนักเรียนในแต่ละข้อ ค่าความเที่ยงที่ได้จากการศึกษาในขั้นตอนการตัดสินใจการสรุปอ้างอิง พบว่ามีค่าสัมประสิทธิ์การสรุปอ้างอิงอยู่ระหว่าง 0.80-0.97 โดยขึ้นอยู่กับชุดของข้อสอบและระดับชั้น

Crehan and Rhoton (1998) ได้ศึกษาเชิงเปรียบเทียบเกี่ยวกับวิธีการให้คะแนนสองวิธีในการประเมินการปฏิบัติงาน โดยวิธีการที่หนึ่งจะใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และวิธีที่สองแบบการให้คะแนนเป็นลำดับขั้น (score categories) วัตถุประสงค์เพื่อค้นหาวิธีการในการปรับปรุงความเป็นปรนัย (objectivity) ความเที่ยง (reliability) และประสิทธิผลของการให้คะแนนการประเมินการปฏิบัติงาน การศึกษานี้ใช้โมเดลการตอบ ในการเปรียบเทียบวิธีการให้คะแนน โดยศึกษาจากนักเรียนที่มีความเร็วในการอ่านในระดับ 5 ในโรงเรียนขนาดใหญ่ที่ใช้มีโปรแกรมการประเมินของตนเอง กลุ่มตัวอย่างจะเป็นอาสาสมัคร (นักเรียนที่จบแล้ว) จำนวน 20 คน ซึ่งจะมีหน้าที่เป็นผู้ประเมิน และอีก 200 คนที่มีความเร็วในการอ่านระดับเดียวกัน แบ่งเป็น 5 กลุ่ม กลุ่มละ 40 คน วิธีการดำเนินการจะให้ผู้ประเมินจากแต่ละวิธีในการให้คะแนนจำนวน 2 คน ตรวจให้คะแนน 40 รายงาน จากนั้นตามด้วยการเปรียบเทียบคะแนนทั้งสองสำหรับแต่ละคำตอบ ภายใต้เงื่อนไขของการให้คะแนนของแต่ละวิธีผลการศึกษาพบว่า ไม่พบความแตกต่างเมื่อใช้วิธีการให้คะแนนทั้งสองวิธีการ

Kristies et al. (1998) ได้ศึกษาเกี่ยวกับผลกระทบ จากการใช้วิธีการให้คะแนนที่แตกต่างกันที่มีต่อความเที่ยงของคะแนนจากการประเมินผลการปฏิบัติงาน โดยศึกษาจากนักเรียนทั้งหมด 6 โรงเรียนจำนวน 100 ถึง 200 คน ซึ่งใช้ผลการปฏิบัติงานวิทยาศาสตร์ วิธีการดำเนินจะประเมิน

โดยใช้วิธีการให้คะแนนที่แตกต่างกันสองวิธี ดังนี้ คือ วิธีที่หนึ่งเป็นการให้คะแนนแบบรวม (Focused holistic, F-H) และวิธีที่สอง แบบแยก (Analytic impression, A-I) ผลการศึกษาพบว่า ไม่มีความแตกต่างของทั้งสองวิธีในการตัดสินนักเรียนแต่ละคนหรือกลุ่ม อย่างไรก็ตามถึงแม้ว่าจะไม่มีความแตกต่างกันผู้ประเมินได้ลงความเห็นเห็นว่า วิธีการแบบแยกจะให้ข้อมูลที่เป็นประโยชน์ในการปรับปรุง พัฒนาการสอนได้ดี

Penny et al. (2000) ได้ศึกษาเกี่ยวกับการใช้มาตรวัดระดับ เพื่อใช้ในการขยายความเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ผู้วิจัยได้ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบเกี่ยวกับการเขียน 4 ประเด็น ซึ่งประกอบด้วย มาตรวัดระดับ 4 ระดับในการให้คะแนนโดยนำตัวอย่างผลงานการเขียนจำนวน 120 ผลงานจากโรงเรียนในจอร์เจียเกรด 11 ที่ในการประเมินการเขียน ผู้ประเมินจะต้องให้คะแนนผลงาน ซึ่งจะต้องทำเครื่องหมาย + หรือ - ผลการศึกษาพบว่า การขยายความในมาตรวัดระดับนั้นเป็นตัวบ่งชี้เพิ่มค่าความเที่ยงของผู้ประเมินได้ และพบว่า ในการขยายแต่ละระดับคะแนน ครูผู้สอนสามารถนำข้อมูลนั้นมาใช้ในการวินิจฉัยเกี่ยวกับการประเมินการปฏิบัติของนักเรียนได้

งานวิจัยในประเทศ

งานวิจัยเกี่ยวกับเกณฑ์การให้คะแนนและทฤษฎีการสรุปอ้างอิงยังมีน้อยมาก ผู้วิจัยได้ศึกษางานวิจัยสรุปได้ว่า งานวิจัยส่วนมากจะศึกษาเปรียบเทียบวิธีการตรวจ จำนวนผู้ตรวจที่แตกต่างกัน ซึ่งบางงานวิจัยได้นำทฤษฎีการสรุปอ้างอิงมาประยุกต์ใช้ในการวิจัย ผู้วิจัยจึงได้รวบรวมไว้ดังนี้

ไพรัตน์ วงษ์นาม (2533) ศึกษาการประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงในแบบทดสอบความเรียงและเปรียบเทียบผลของวิธีตรวจ การชี้แจง และการรู้ผลการเรียนของผู้สอบ โดยวัดความสามารถในการแสดงความคิดเห็น เกี่ยวกับข่าวและเหตุการณ์ ของนักเรียนชั้นประถมศึกษาปีที่ 6 พร้อมทั้งเลือกวิธีตรวจ จำนวนผู้ตรวจ และจำนวนข้อสอบที่ให้ค่าสัมประสิทธิ์การสรุปอ้างอิง .50 เป็นอย่างต่ำ กลุ่มตัวอย่างที่ใช้เป็น นักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 3 คน ข้อสอบแบบความเรียงจำนวน 5 ข้อ และครูผู้ตรวจจำนวน 20 คน แบ่งเป็น 2 กลุ่ม กลุ่มละ 10 คน กลุ่มหนึ่งตรวจแบบประเมินรวม และอีกกลุ่มตรวจแบบวิธีวิเคราะห์ ในแต่ละวิธีทำการตรวจ 3 ครั้ง ดังนี้ ตรวจโดยใช้ประสบการณ์เดิมของผู้ตรวจให้คะแนน ตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยโดยไม่รู้ผลการเรียนของผู้ตอบ และตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยโดยรู้ผลการเรียนของผู้ตอบ ผลการวิจัยพบว่า เมื่ออ้างอิงสรุปไปยังเอกภาพของข้อสอบและผู้ตรวจพร้อมกัน (M1) ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบความเรียง ที่ตรวจโดยวิธีประเมินผลรวม มีค่าระหว่าง 0.3328 - 0.4782 และตรวจโดยวิธีวิเคราะห์ มีค่าอยู่ระหว่าง 0.3348 - 0.5895 เมื่ออ้างอิง

สรุปไปยังเอกภพของผู้ตรวจอย่างเดียว (M2) ค่าสัมประสิทธิ์การสรุปอ้างอิงของวิธีประเมินรวมมีค่าอยู่ระหว่าง 0.5985 – 0.7761 วิธีตรวจการชี้แจง และการรู้ผลการเรียนของผู้ตอบต่างกัน ไม่มีผลต่อความแตกต่างของค่าสัมประสิทธิ์การสรุปอ้างอิงสำหรับ M1 แต่ใน M2 ค่าสัมประสิทธิ์การสรุปอ้างอิงของการตรวจตามที่ได้รับการชี้แจงมีค่าสูงกว่าการตรวจโดยใช้ประสบการณ์เดิมเฉพาะกรณีที่ผู้ตรวจได้รับการฝึกอบรม ไม่รู้ผลการเรียนของผู้ตอบเท่านั้น ถ้ารู้ผลการเรียนของผู้ตอบ ค่าสัมประสิทธิ์การสรุปอ้างอิงไม่แตกต่างกัน ค่าสัมประสิทธิ์การสรุปอ้างอิงของการตรวจไม่รู้ประวัติการเรียนของผู้ตอบ มีค่าสูงกว่าผู้ตรวจรู้ประวัติการเรียนของผู้ตอบ วิธีตรวจ จำนวนข้อสอบ และจำนวนผู้ตรวจที่ให้ค่าสัมประสิทธิ์การสรุปอ้างอิงไม่น้อยกว่า 0.5 คือวิธีตรวจแบบวิเคราะห์ตามที่ได้รับการชี้แจง และผู้ตรวจไม่รู้ผลการเรียนของผู้ตอบ ใช้ข้อสอบ 6 ข้อ ผู้ตรวจอย่างน้อย 5 คน สำหรับ M1 และผู้ตรวจ 5 คน ข้อสอบ 5 ข้อ สำหรับ M2

จิราพร อัครสมพงษ์ (2536) ได้ศึกษาเปรียบเทียบความสอดคล้องของคะแนนแบบทดสอบความเรียงจากผู้ตรวจให้คะแนน และวิธีการตรวจต่างกัน วัตถุประสงค์เพื่อศึกษาเปรียบเทียบความสอดคล้องในการให้คะแนนของผู้ตรวจให้คะแนน 2 คน ที่ตรวจแบบทดสอบความเรียงโดยใช้วิธีการเดียวกัน และความสอดคล้องในการให้คะแนนระหว่างวิธีการตรวจที่แตกต่าง พบว่าความสัมพันธ์ในการให้คะแนนระหว่างผู้ตรวจให้คะแนนด้วยวิธีเดียวกัน ผู้ตรวจให้คะแนนที่ตรวจด้วยวิธีเดียวกันมีความสัมพันธ์สอดคล้องกัน ในทิศทางเดียวกันและมีค่าสัมประสิทธิ์สหสัมพันธ์ของการตรวจแบบวิธีวิเคราะห์ให้เสร็จรายข้อ กับการตรวจแบบวิธีการประเมินรวมให้เสร็จเป็นรายข้อมีค่าสูงสุด คือ 0.8116

อุษณีย์ บัวศิริพันธุ์ (2543) ศึกษาเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจให้คะแนน และประสบการณ์ของผู้ตรวจให้คะแนนแตกต่างกัน ผลการศึกษาเกี่ยวกับวิธีการตรวจให้คะแนน การตรวจแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนนของผู้ตรวจคะแนนเอง เมื่อมีจำนวนผู้ตรวจให้คะแนน 3, 4, 5 และ 6 คน มีค่าเฉลี่ยอยู่ระหว่าง 8.79 – 9.05 โดยการตรวจให้คะแนนเมื่อมีจำนวนผู้ตรวจให้คะแนนจำนวน 4 คน คะแนนเฉลี่ยมีแนวโน้มสูงที่สุด รองลงมา คือ มีจำนวนผู้ตรวจให้คะแนน 5 คน และเมื่อมีจำนวนผู้ตรวจให้คะแนน 6 คน คะแนนเฉลี่ยมีแนวโน้มต่ำที่สุด ส่วนการตรวจให้คะแนนแบบภาพรวมโดยใช้เกณฑ์การให้คะแนน เมื่อจำนวนผู้ตรวจให้คะแนน 3, 4, 5 และ 6 คะแนนเฉลี่ยมีแนวโน้มสูงที่สุด รองลงมาคือ จำนวนผู้ตรวจให้คะแนน 3 และ 6 คน และเมื่อมีจำนวนผู้ตรวจให้คะแนน 5 คน คะแนนเฉลี่ยมีแนวโน้มต่ำที่สุด

พรธณี เจียมสุขบุตร (2543) ได้ศึกษาเปรียบเทียบความเที่ยงของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจ และวิธีการตรวจต่างกัน การศึกษาค้นคว้าครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าความเที่ยงของแบบทดสอบวัดความสามารถใน

การแก้ไขข้อปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจ 2 คน และ 3 คน โดยใช้วิธีการตรวจวิธี ประเมินรวม และวิธีวิเคราะห์ย่อย ศึกษาจากกลุ่มตัวอย่างนักเรียนชั้นมัธยมศึกษาปีที่ 2 จำนวน 70 คน ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัด กรุงเทพฯ ผลการศึกษาพบว่า 1) ค่าความเที่ยงของแบบทดสอบโดยการตรวจให้คะแนนด้วยวิธีวิเคราะห์ที่มีจำนวนผู้ตรวจ 2 คน และ 3 คน มีค่าเท่ากับ 0.92 และ 0.96 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ส่วนการตรวจให้คะแนนด้วยวิธีรวมที่มีจำนวนผู้ตรวจ 2 และ 3 คน มีค่าเท่ากับ 0.97 และ 0.98 ตามลำดับ ซึ่งไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ 2) ค่าความเที่ยงของแบบทดสอบที่ใช้เกณฑ์การให้คะแนนแบบวิธีรวมและวิธีวิเคราะห์ ของผู้ตรวจให้คะแนน 2 คน มีค่าเท่ากับ 0.97 และ 0.92 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนแบบทดสอบที่ใช้วิธีการตรวจให้คะแนนวิธีประเมินรวม และวิธีวิเคราะห์ ของผู้ตรวจให้คะแนน 3 คน เท่ากับ 0.98 และ 0.96 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

จากการสังเคราะห์งานวิจัยที่เกี่ยวข้องกับการศึกษาเกณฑ์การให้คะแนนข้างต้น ผู้วิจัยจึงสนใจศึกษาในเปรียบเทียบค่าความเที่ยงโดยใช้ทฤษฎีการสรุปอ้างอิงเพื่อหาจำนวนผู้ประเมิน และจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนที่ต่างกัน เพราะจากการศึกษางานวิจัยที่นำทฤษฎีการสรุปอ้างอิง ทฤษฎีนี้สามารถบอกระดับความน่าเชื่อถือของจำนวนผู้ประเมินได้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนจากผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) และเพื่อหาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) ในบทนี้ผู้วิจัยได้นำเสนอลำดับขั้นตอนในการวิจัยดังนี้

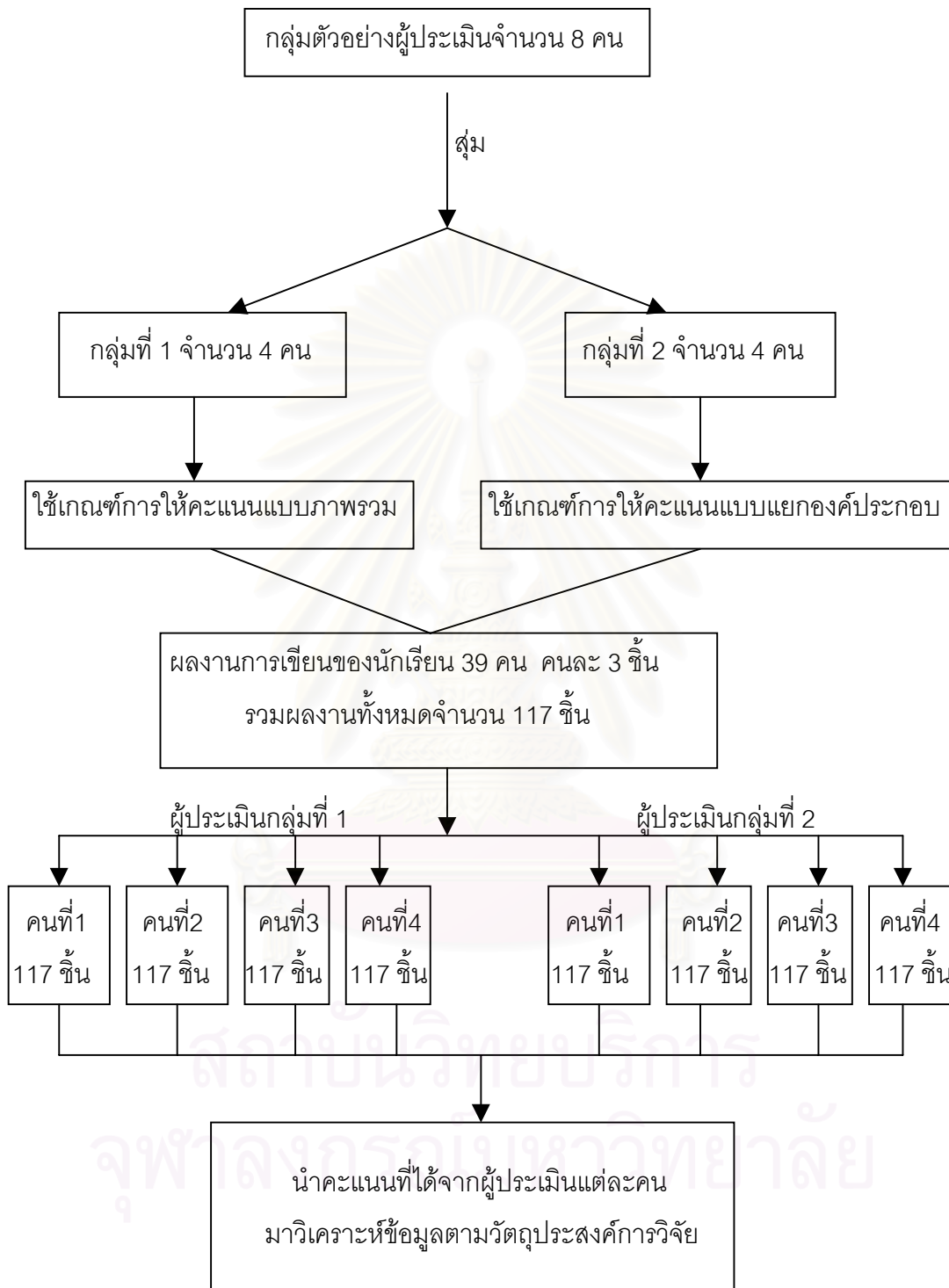
ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการศึกษาครั้งนี้เป็นเป็นผู้ประเมินซึ่งเป็นครูผู้สอนวิชาภาษาไทย ชั้นประถมศึกษาปีที่ 5-6 ในโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาเอกชนใน กรุงเทพมหานครและปริมณฑล และผลงานการเขียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2544 ของโรงเรียนที่ผู้วิจัยคัดเลือกซึ่งเป็นโรงเรียนที่ยินดีให้ความร่วมมือในเขตกรุงเทพมหานคร จำนวน 1 โรงเรียน ได้แก่ โรงเรียนไผ่ทอคมศึกษา

กลุ่มตัวอย่าง ในการวิจัยครั้งนี้มีกลุ่มตัวอย่างที่ศึกษาประกอบด้วย

1. กลุ่มตัวอย่าง ผลงานการเขียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2544 จำนวน 1 ห้องเรียน ของโรงเรียนไผ่ทอคมศึกษา
2. กลุ่มตัวอย่างผู้ประเมิน เนื่องจากในงานวิจัยครั้งนี้ต้องใช้ครูผู้สอนวิชาภาษาไทย ชั้นประถมศึกษาปีที่ 5-6 เป็นผู้ประเมินผลงานการเขียนซึ่งต้องใช้เวลาและความสนใจ ตั้งใจในการประเมินผลงานการเขียน ผู้วิจัยทำการสุ่มครูผู้ประเมินโดยวิธีการแบบเจาะจง (purposive sampling) เพราะต้องหาผู้ประเมินที่เต็มใจและมีเวลาช่วยประเมินอย่างเต็มที่ มีประสบการณ์ ในการสอนอย่างน้อย 5 ปี ครูที่ได้รับเลือกเป็นกลุ่มตัวอย่างเป็นครูจาก โรงเรียนไผ่ทอคมศึกษา จำนวน 3 คน โรงเรียนอัสสัมชัญแผนกประถม จำนวน 3 คน และ โรงเรียนศิริวิทยา จำนวน 2 คน รวมจำนวนทั้งหมด 8 คน แบ่งครูออกเป็น 2 กลุ่มๆละ 4 คนโดยใช้วิธีสุ่มอย่างง่ายโดยการ จับสลาก ซึ่งผู้ประเมินกลุ่มที่ 1 จะใช้เกณฑ์การให้คะแนนแบบภาพรวม และ ผู้ประเมินกลุ่มที่ 2 จะใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบในการให้คะแนนผลงานการเขียน โดยผู้ประเมินแต่ละคนจะตรวจผลงานของนักเรียนจำนวน 39 คนซึ่งนักเรียนแต่ละคนมีผลงานคนละ 3 ชิ้น

ดังนั้น ผู้ประเมินแต่ละคนจะตรวจผลงานรวมทั้งหมด 117 ชิ้น ผู้วิจัยได้สรุปขั้นตอนดังแผนภาพที่ 2



แผนภาพที่ 2 แสดงขั้นตอนการตรวจผลงานของผู้ประเมิน

เครื่องมือในการดำเนินการวิจัย

เครื่องมือในการดำเนินการวิจัยประกอบด้วย

- 1) เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric)
- 2) เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)
- 3) แบบบันทึกคะแนน

การสร้างและพัฒนาเครื่องมือ

การสร้างและพัฒนาเครื่องมือ ได้แบ่งออกเป็น 2 ส่วนใหญ่ ดังต่อไปนี้

- 1) การสร้างและพัฒนาเกณฑ์การให้คะแนน
- 2) การสร้างแบบบันทึกคะแนน

การสร้างและพัฒนาเกณฑ์การให้คะแนน

1. ศึกษาหลักสูตรประถมศึกษา คู่มือการประเมินผลการเรียนระดับประถมศึกษาแบบเรียนที่ใช้คู่มือครู และแผนการสอนของครูรวมทั้งจุดมุ่งหมายของหลักสูตร จุดประสงค์กลุ่ม ประสพการณ์และจุดประสงค์รายวิชาในวิชาภาษาไทย ตลอดจนเนื้อหาในแบบเรียน ศึกษาเอกสารรายงานวิจัยเกี่ยวกับการประเมินตามสภาพจริง และเกณฑ์การให้คะแนน เพื่อเป็นแนวทางในการสร้างเกณฑ์การให้คะแนนผลงานที่สอดคล้องกับวัตถุประสงค์ผลงานการเขียน

2. กำหนดวิชาและระดับชั้นที่จะนำผลงานการเขียนมาทำการตัดสินให้คะแนนวิชาที่ใช้ในการวิจัย คือ วิชาภาษาไทย ทั้งนี้เพราะจากงานวิจัยของ Waltman (1998) เป็นการศึกษา การตรวจผลงานการประเมินการปฏิบัติการเขียนโดยใช้วิธีการให้คะแนนทั้งสองแบบได้ผลว่า เกณฑ์การให้คะแนนทั้งสองรูปแบบให้ผลที่แตกต่างกัน และพบว่าวิธีการให้คะแนนแบบแยก องค์ประกอบ จะใช้เวลาในการตรวจมาก แต่ให้ค่าความสอดคล้องภายในของผู้ตรวจสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม จึงนำมาทดลองใช้ตรวจผลงานการเขียนวิชาภาษาไทย ส่วนสาเหตุ ที่เลือกชั้นประถมศึกษาปีที่ 5 เนื่องจากในระดับชั้นประถมศึกษาปีที่ 6 นั้นส่วนมากเนื้อหาจะเป็นการทบทวนเนื้อหาที่สำคัญ และมีกิจกรรมที่มุ่งการศึกษาต่อเป็นสำคัญ

3. ผู้วิจัยนำผลงานที่เกี่ยวกับทักษะทางการเขียนในแฟ้มสะสมงานวิชาภาษาไทยที่มี 3 ลักษณะ คือ การเขียนเรื่องประกอบภาพ การเขียนจดหมาย และการบันทึกความรู้เรื่องที่อ่าน ของนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนที่เป็นกลุ่มตัวอย่างมาทำการศึกษาวัตถุประสงค์ของกิจกรรมนั้นๆ ซึ่งพบว่ามี 3 วัตถุประสงค์ ดังนี้

1. เขียนตามรูปแบบที่กำหนด โดยลำดับความคิดเหตุการณ์อย่างต่อเนื่อง
2. เขียนแสดงความคิดเห็นอย่างเสรีและเชิงสร้างสรรค์
3. เขียนบันทึกความรู้เรื่องที่อ่านได้อย่างถูกต้อง

4. จากนั้นผู้วิจัยสร้างและพัฒนาเกณฑ์การให้คะแนนภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบโดยมีขั้นตอนดังนี้

4.1 เกณฑ์การให้คะแนนแบบภาพรวม ในการสร้างเกณฑ์การให้คะแนนนี้ผู้วิจัยได้ศึกษาเอกสาร และงานวิจัยเกี่ยวกับเกณฑ์การให้คะแนน โดยปรับมาจากเกณฑ์การให้คะแนนการเขียนความเรียงของ เอช แอล จากอบส์ และคณะ (H.L. Jacobs et al. 1981) ซึ่งเขาได้เสนอแนวทางในการตรวจความเรียงโดยประเมินด้วยเกณฑ์ที่เรียกว่า ESL composition profile เป็นเกณฑ์การประเมินผลความสามารถทางการเขียนความเรียงที่ชัดเจน ครอบคลุม องค์ประกอบของการเขียน และแยกประเด็นในการประเมินผลอย่างชัดเจน ทำให้สะดวกและพิจารณางานเขียน ได้ง่าย โดยมีรายละเอียดขององค์ประกอบดังนี้ 1) องค์ประกอบด้านเนื้อหา 2) องค์ประกอบด้านการลำดับเรื่อง 3) องค์ประกอบด้านการใช้ถ้อยคำ 4) องค์ประกอบด้านการใช้ไวยากรณ์ และ 5) องค์ประกอบด้านกลไกการเขียน แต่เนื่องจากผลงานการเขียนอยู่ในระดับชั้นประถมศึกษาปีที่ 5 ผู้วิจัยจึงต้องกำหนดประเด็นหรือองค์ประกอบให้สอดคล้องเหมาะสมกับวัตถุประสงค์ เนื้อหา และระดับชั้นที่ต้องการประเมิน และจากผลงานการวิจัยของ ไพรัตน์ วงษ์นาม (2533) พบว่า ประเด็นการตรวจบางด้านซึ่งกำหนดไว้ใน การตรวจความเรียง เช่น ด้านการใช้ถ้อยคำ ด้านกลไกการเขียน เป็นความสามารถทางภาษาพิเศษ รวมทั้งการกำหนดประเด็นย่อยมากเกินไป จะทำให้ผู้ประเมินตัดสินใจลำบากโดยเกณฑ์การให้คะแนนแบบภาพรวมผู้วิจัยได้อธิบายลักษณะของผลงานใน แต่ละระดับเป็นความเรียงต่อ มีรายละเอียดดังนี้

4.1.1 เกณฑ์การให้คะแนนแบบภาพรวมของการเขียนเรื่องประกอบภาพ จะพิจารณาลักษณะของผลงานซึ่งแยกเป็น 4 ระดับคะแนน คือ ระดับคะแนน 4 หมายถึง ดีมาก ระดับคะแนน 3 หมายถึง ดี ระดับคะแนน 2 หมายถึง พอใช้ และระดับคะแนน 1 หมายถึง ควรปรับปรุง ในประเด็นต่างๆ ดังต่อไปนี้

- ด้านเนื้อหา
- การลำดับเรื่อง
- การใช้ภาษา
- กลไกการเขียน

4.1.2 เกณฑ์การให้คะแนนแบบภาพรวม ของการเขียนจดหมายจะพิจารณา ลักษณะของผลงานซึ่งแยกเป็น 4 ระดับคะแนน คือ ระดับคะแนน 4 หมายถึง ดีมาก ระดับคะแนน 3 หมายถึง ดี ระดับคะแนน 2 หมายถึง พอใช้ และระดับคะแนน 1 หมายถึง ควรปรับปรุง ใน ประเด็นต่างๆ ดังต่อไปนี้

- แบบฟอร์มของการเขียนจดหมาย
- การลำดับเรื่อง
- การใช้ภาษา
- กลไกการเขียน

4.1.3 เกณฑ์การให้คะแนนแบบภาพรวมของการบันทึกความรู้เรื่องี่อ่านจะ พิจารณาลักษณะของผลงานซึ่งแยกเป็น 4 ระดับคะแนน คือ ระดับคะแนน 4 หมายถึง ดีมาก ระดับ คะแนน 3 หมายถึง ดี ระดับคะแนน 2 หมายถึง พอใช้ และระดับคะแนน 1 หมายถึง ควรปรับ ปรุง ในประเด็นต่างๆ ดังต่อไปนี้

- การระบุประเด็นสำคัญ
- การลำดับเรื่อง
- การใช้ภาษา
- กลไกการเขียน

4.2 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ในการสร้างเกณฑ์การให้คะแนน นี้ผู้วิจัยใช้หลักการสร้างเช่นเดียวกับการสร้างเกณฑ์การให้คะแนนแบบภาพรวม แต่ในขั้นตอนนี้ ผู้ วิจัยได้แยกรายการหรือองค์ประกอบที่จะพิจารณาให้ระดับคะแนนผลงานเป็นมิติต่างๆ ใน รูปแบบตาราง 2 มิติ ซึ่งแต่ละองค์ประกอบจะแบ่งเป็นระดับคุณภาพของผลงาน 4 ระดับ และ ได้ กำหนดคะแนนเป็นช่วง

4.2.1 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบของการเขียนเรื่อง ประกอบภาพจะพิจารณาลักษณะของผลงานซึ่งแยกเป็น 4 ระดับคะแนน คือ ระดับคะแนน 4 (ช่วง คะแนน 14-16) หมายถึง ดีมาก ระดับคะแนน 3 (ช่วงคะแนน 11-13) หมายถึง ดี ระดับคะแนน 2 (ช่วงคะแนน 7-10) หมายถึง พอใช้ และระดับคะแนน 1 (ช่วงคะแนน 4-6) หมายถึง ควรปรับปรุง ในประเด็นต่างๆ ดังต่อไปนี้

- เนื้อหา
- การลำดับเรื่อง
- การใช้ภาษา
- กลไกการเขียน

4.2.2 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบของการเขียนจดหมายจะพิจารณาลักษณะของผลงานซึ่งแยกเป็น 4 ระดับคะแนน คือ ระดับคะแนน 4 หมายถึง ดีมาก ระดับคะแนน 3 หมายถึง ดี ระดับคะแนน 2 หมายถึง พอใช้ และระดับคะแนน 1 หมายถึง ควรปรับปรุง ในประเด็นต่างๆ ดังต่อไปนี้

- แบบฟอร์มของจดหมาย
- เนื้อหา
- การใช้ภาษา
- กลไกการเขียน

4.2.3 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบของการบันทึกความรู้เรื่อง ที่อ่านจะพิจารณาลักษณะของผลงานซึ่งแยกเป็น 4 ระดับคะแนน คือ ระดับคะแนน 4 หมายถึง ดีมาก ระดับคะแนน 3 หมายถึง ดี ระดับคะแนน 2 หมายถึง พอใช้ และระดับคะแนน 1 หมายถึง ควรปรับปรุง ในประเด็นต่างๆ ดังต่อไปนี้

- การระบุประเด็นสำคัญ
- การลำดับเรื่อง
- การใช้ภาษา
- กลไกการเขียน

5. ผู้วิจัยนำเกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ผู้ทรงคุณวุฒิทางด้านภาษาไทย จำนวน 5 ท่านซึ่งเกณฑ์ที่ใช้ในการคัดเลือก ผู้ทรงคุณวุฒิ คือต้องเป็นผู้ที่มีประสบการณ์การสอนภาษาไทยไม่น้อยกว่า 10 ปี และมีผลงานดีเด่นทางด้านวิชาภาษาไทยพิจารณาความเหมาะสมและให้ข้อคิดเห็นเกี่ยวกับเกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบที่สร้างขึ้นปรากฏว่าผู้ทรงคุณวุฒิให้ปรับปรุงแก้ไขภาษาให้มีความชัดเจนขึ้น ปรับปรุงประเด็นย่อยๆของแต่ละองค์ประกอบ ซึ่งผู้ทรงคุณวุฒิให้เพิ่มน้ำหนักเป็นเปอร์เซ็นต์ นอกจากนี้ผู้ทรงคุณวุฒิให้ข้อเสนอแนะว่าเกณฑ์การให้คะแนนนั้นต้องพิจารณาให้เหมาะสมกับระดับชั้นที่จะนำไปประเมิน

6. ผู้วิจัยปรับปรุงตามผู้ทรงคุณวุฒิในด้านภาษาให้ชัดเจนและปรับปรุงประเด็นย่อยๆขององค์ประกอบต่างๆ โดยในองค์ประกอบด้านกลไกการเขียนและองค์ประกอบด้านเนื้อหาแล้วให้อาจารย์ที่ปรึกษาตรวจสอบอีกครั้งจนได้เกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบที่สมบูรณ์

ตัวอย่างเกณฑ์การให้คะแนน

1. เกณฑ์การให้คะแนนแบบภาพรวม

ตัวอย่าง เกณฑ์การให้คะแนนแบบภาพรวม

(หมายเหตุ การให้คะแนนผลงาน ให้พิจารณาจากคุณลักษณะที่สำคัญที่สุดโดยรวมของผลงาน)

การเขียนเรื่องประกอบภาพ

| ระดับ คะแนน | ลักษณะของผลงาน |
|--------------------|---|
| 4 (ดีมาก) | ผลงานมีลักษณะตามเกณฑ์ดังนี้ - มีความรู้ในเรื่องที่เขียนดีมาก เขียนได้ใจความน่าสนใจ มีความคิดสร้างสรรค์ดีมาก - เนื้อหามีความสัมพันธ์กับภาพประกอบและชื่อเรื่อง - เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความ กะทัดรัด และมีความต่อเนื่องดีมาก - ใช้ถ้อยคำ สำนวนภาษาได้อย่างถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อความหมายได้ชัดเจนดีมาก - ไม่มีข้อผิดพลาดเลยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ |
| 3 (ดี) | ผลงานมีลักษณะตามเกณฑ์ดังนี้ - มีความรู้ในเรื่องที่เขียนดี เขียนได้ใจความน่าสนใจ มีความคิดสร้างสรรค์ดี เนื้อหาส่วนใหญ่มีความสัมพันธ์กับภาพประกอบและชื่อเรื่อง - เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความ กะทัดรัด และมีความต่อเนื่องดี - มีข้อผิดพลาดเล็กน้อยในการใช้ถ้อยคำ สำนวนภาษา สามารถสื่อความหมายได้ไม่ทำให้ความหมายเปลี่ยน - มีข้อผิดพลาดเล็กน้อยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 10 % แต่ไม่เกิน 50 % |
| 2 (พอใช้) | ผลงานมีลักษณะตามเกณฑ์ดังนี้ - มีความรู้ในเรื่องที่เขียนพอสมควร เขียนได้ใจความพอสมควร ขาดความคิดสร้างสรรค์ เนื้อหาบางส่วนไม่สัมพันธ์กับภาพประกอบและชื่อเรื่อง - เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความ กะทัดรัด และมีความต่อเนื่องค่อนข้างน้อย - ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้อง และไม่เหมาะสม บางครั้งไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดค่อนข้างมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | ผลงานมีลักษณะตามเกณฑ์ดังนี้ - มีความรู้ในเรื่องที่เขียนน้อย เขียนเนื้อหาไม่ตรงกับชื่อเรื่อง ไม่น่าสนใจ ไม่ได้ใจความ ขาดความคิดสร้างสรรค์ ส่วนภาพประกอบและชื่อเรื่องไม่สัมพันธ์กับเนื้อหา - ไม่สามารถเรียบเรียงความคิดอย่างสมเหตุสมผล และไม่มีมีความต่อเนื่อง - ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้องและไม่เหมาะสม ไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 90 % |

2. เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

ตัวอย่าง เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ การเขียนเรื่องประกอบภาพ

| ระดับ คะแนน | เนื้อหา | การลำดับเรื่อง | การใช้ภาษา | กลไกการเขียน |
|--------------------|---|--|---|--|
| 4 (ดีมาก) | มีความรู้ในเรื่องที่เขียนดี มาก เขียนได้ใจความ น่าสนใจ มีความคิดสร้าง สรรค์ดีมาก เนื้อหามี ความสัมพันธ์กับภาพ ประกอบและชื่อเรื่อง | เรียบเรียงความคิด อย่างสมเหตุสมผล เขียนได้ความกะทัดรัด และมีความต่อเนื่องดี มาก | ใช้ถ้อยคำ สำนวนภาษาได้ อย่างถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อ ความหมายได้ชัดเจนดีมาก | ไม่มีข้อผิดพลาดเลยใน การสะกดคำและ การใช้เครื่องหมายวรรค ตอนต่างๆ |
| 3 (ดี) | มีความรู้ในเรื่องที่เขียนดี เขียนได้ใจความ น่าสนใจ มีความคิดสร้างสรรค์ดี เนื้อหาส่วนใหญ่มีความ สัมพันธ์กับภาพประกอบ และชื่อเรื่อง | เรียบเรียงความคิด อย่าง สมเหตุสมผล เขียนได้ความกะทัดรัด และมีความต่อเนื่องดี | มีข้อผิดพลาดเล็กน้อย ในการใช้ถ้อยคำ สำนวน ภาษา สามารถสื่อความ หมายได้ไม่ทำให้ความ หมายเปลี่ยน | มีข้อผิดพลาดเล็กน้อยใน การสะกดคำและ การใช้ เครื่องหมายวรรคตอน ต่างๆ เกิน 10 % แต่ ไม่เกิน 50 % |
| 2 (พอใช้) | มีความรู้ในเรื่องที่เขียน พอสมควร เขียนได้ ใจความพอสมควร ขาดความคิดสร้างสรรค์ เนื้อหามีบางส่วน ไม่สัมพันธ์กับภาพ ประกอบและชื่อเรื่อง | เรียบเรียงความคิด อย่างสมเหตุสมผล เขียนได้ความกะทัดรัด และมีความต่อเนื่อง ค่อนข้างน้อย | ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้องและ ไม่เหมาะสม บางครั้ง ไม่สามารถสื่อความหมาย ให้เข้าใจได้ | มีข้อผิดพลาดค่อนข้าง มากในการสะกดคำและ การใช้เครื่องหมายวรรค ตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | มีความรู้ในเรื่องที่เขียน น้อย เขียนเนื้อหาไม่ตรง กับชื่อเรื่อง ไม่น่าสนใจ ไม่ได้ใจความ ขาดความ คิดสร้างสรรค์ ส่วนภาพ ประกอบและชื่อเรื่อง ไม่สัมพันธ์กับเนื้อหา | ไม่สามารถเรียบเรียง ความคิดอย่างสมเหตุ สมผลและไม่มีความ ต่อเนื่อง | ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้องและไม่เหมาะสม ไม่สามารถสื่อความหมาย ให้เข้าใจได้ | มีข้อผิดพลาดมากในการ สะกดคำและการใช้เครื่อง หมายวรรคตอนต่างๆ เกิน 90 % |

การสร้างแบบบันทึกคะแนน

1. ผู้วิจัยสร้างแบบบันทึกคะแนนของเกณฑ์การให้คะแนนแบบภาพรวม ดังตัวอย่างต่อไปนี้

ตัวอย่าง แบบบันทึกการให้คะแนนผลงานแบบภาพรวม

คำชี้แจง ให้ผู้ประเมินใช้เกณฑ์การให้คะแนนผลงานการเขียนเรื่องประกอบภาพแบบภาพรวมพิจารณาผลงานในแฟ้มสะสมงาน แล้วตัดสินผลงานตามคุณภาพของรายการประเมินโดยใส่เครื่องหมาย ✓ ลงในช่องตาราง

| ระดับคะแนน | คุณภาพของผลงานโดยรวม |
|------------|----------------------|
| 4 | ✓ |
| 3 | |
| 2 | |
| 1 | |

ได้คะแนน

4

เกณฑ์การประเมิน

คะแนน 4 หมายถึง ดีมาก

คะแนน 3 หมายถึง ดี

คะแนน 2 หมายถึง พอใช้

คะแนน 1 หมายถึง ควรปรับปรุง

2. ผู้วิจัยสร้างแบบบันทึกคะแนนของเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ดังตัวอย่างต่อไปนี้

ตัวอย่าง แบบบันทึกการให้คะแนนผลงานแบบแยกองค์ประกอบ

คำชี้แจง ให้ผู้ประเมินใช้เกณฑ์การให้คะแนนผลงานการเขียนเรื่องประกอบภาพแบบแยกองค์ประกอบพิจารณาผลงานในแฟ้มสะสมงานแล้วตัดสินผลงานตามคุณภาพของรายการประเมินโดยใส่เครื่องหมาย ✓ ลงในช่องตาราง

| ระดับคะแนน | รายการประเมิน | | | |
|------------|---------------|------------------------|------------|--------------|
| | เนื้อหา | การเรียบเรียงเรื่องราว | การใช้ภาษา | กลไกการเขียน |
| 4 | ✓ | | | ✓ |
| 3 | | ✓ | ✓ | |
| 2 | | | | |
| 1 | | | | |

ได้คะแนนรวม

14

คะแนนของผลงาน เท่ากับ.....4.....

เกณฑ์การประเมิน

| ช่วงคะแนน | ระดับคุณภาพ |
|-----------|-----------------------|
| 14 – 16 | 4 หมายถึง ดีมาก |
| 11 – 13 | 3 หมายถึง ดี |
| 7 – 10 | 2 หมายถึง พอใช้ |
| 4 – 6 | 1 หมายถึง ควรปรับปรุง |

3. นำแบบบันทึกการให้คะแนนแบบภาพรวมและแบบแยกองค์ประกอบ ไปให้อาจารย์ที่ปรึกษาตรวจสอบ และปรับปรุงแก้ไข ตามคำแนะนำ จนได้แบบบันทึกการให้คะแนนแบบภาพรวมและแบบแยกองค์ประกอบที่สมบูรณ์

การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลดำเนินการดังนี้

1. นำหนังสือขอความร่วมมือจากบัณฑิตวิทยาลัย ขออนุญาตเก็บข้อมูลในโรงเรียนที่ผู้วิจัยคัดเลือกเป็นกลุ่มตัวอย่าง
2. ดำเนินการวิจัยโดยผู้วิจัยขอความร่วมมือในการเก็บรวบรวมแฟ้มสะสมงานในวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2544 จำนวน 1 ห้องเรียน
3. ติดต่อประสานงานและขออนุญาตผู้บริหารโรงเรียนและครูผู้ประเมินที่ใช้เป็นกลุ่มตัวอย่างเพื่อกำหนด วัน เวลา ในการประเมิน
4. ผู้วิจัยนำผลงานของนักเรียนในแฟ้มสะสมงานจำนวน 39 คน ซึ่งนักเรียนหนึ่งคน มีผลงานในแฟ้มสะสมงานจำนวน 3 ผลงาน รวมทั้งหมด 117 ผลงาน นำผลงานของนักเรียน มาใส่เลขที่ใหม่ แล้วถ่ายเอกสารเป็น 8 ชุด เพื่อจัดให้แก่ผู้ประเมิน ดำเนินการให้คะแนนดังนี้
 - 4.1 จัดผลงานแฟ้มสะสมงาน 8 ชุด สำหรับผู้ประเมิน 8 คน โดยผู้ประเมินกลุ่มที่ 1 จำนวน 4 คน ที่ใช้เกณฑ์การให้คะแนนแบบภาพรวมซึ่งผู้ประเมินแต่ละคนจะได้ผลงานของ นักเรียนจำนวน 117 ผลงาน เกณฑ์การให้คะแนนแบบภาพรวมจำนวน 1 ชุด และแบบบันทึกคะแนนแบบภาพรวมจำนวน 117 ใบ ส่วนผู้ประเมินกลุ่มที่ 2 จำนวน 4 คน ที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบซึ่งผู้ประเมินแต่ละคนจะได้ผลงานของนักเรียนจำนวน 117 ผลงาน เกณฑ์การให้คะแนนแบบแยกองค์ประกอบจำนวน 1 ชุด และแบบบันทึกคะแนนแบบแยกองค์ประกอบจำนวน 117 ใบ
 - 4.2 จากนั้นผู้วิจัยชี้แจงให้ผู้ประเมินในกลุ่มที่ 1 ทราบถึงรายละเอียดของการตรวจ โดยใช้เกณฑ์การให้คะแนนแบบภาพรวม มีวิธีดังนี้

1. อ่านรายละเอียดของเกณฑ์การให้คะแนนในประเด็นต่างๆ ทั้ง 4 ระดับ
คุณภาพให้เข้าใจก่อนจึงเริ่มทำการประเมิน
2. อ่านผลงานของนักเรียนแล้วตรวจให้คะแนนตามเกณฑ์การให้คะแนนของแต่ละผลงานโดยตรวจทีละข้อ จนครบทุกคน
3. กรอคะแนนลงในแบบบันทึกคะแนนแบบภาพรวมจนครบทั้ง 117 ผลงาน
สำหรับผู้ประเมินกลุ่มที่ 2 ที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีวิธีการตรวจดังนี้
 1. อ่านเกณฑ์การให้คะแนนในแต่ละองค์ประกอบที่แยกเป็นประเด็นต่างๆ ทั้ง 4 ระดับคุณภาพให้เข้าใจก่อนจึงเริ่มทำการประเมิน
 2. อ่านผลงานของนักเรียนแล้วตรวจให้คะแนนตามเกณฑ์การให้คะแนนของแต่ละผลงานโดยตรวจทีละข้อ จนครบทุกคน
 3. ตรวจให้คะแนนผลงานโดยให้คะแนนในแต่ละองค์ประกอบย่อยนั้น ตามระดับคุณภาพแล้วค่อยรวมเป็นคะแนนรวมทั้งผลงาน กรอคะแนนลงในแบบบันทึกคะแนนแบบแยกองค์ประกอบจนครบทั้ง 117 ผลงาน

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลมีการวิเคราะห์ข้อมูลดังนี้

1. หาค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมินจำนวน 8 คน ได้แก่ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน
 2. หาค่าความเที่ยงของคะแนนการประเมินผลงานการเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนนจากค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) โดยใช้โปรแกรม GENOVA Version 2.2
- การออกแบบการวัด
- 2.1.1 การวัดมี 2 ฟาเซต (facets) ดังนี้
Facet ที่ 1 : ผลงาน (แทนสัญลักษณ์ด้วย I)
Facet ที่ 2 : ผู้ประเมิน (แทนสัญลักษณ์ด้วย R)
 - 2.2.2 รูปแบบการให้คะแนนเป็น PXIXR Cross Design กล่าวคือ ผู้ประเมิน ทุกคนตรวจให้คะแนนผลงานทุกผลงานของนักเรียนทุกคน

การคำนวณ G-STUDY เพื่อประมาณความแปรปรวนจากแหล่งต่างๆ ดังต่อไปนี้

ตาราง 1 การประมาณค่าความแปรปรวนจากแหล่งต่างๆ ของ PXIXR Design
(ศิริชัย กาญจนวาสี, 2542)

| Effect | σ^2 | Expected Mean Squares |
|--------|--|---|
| P | $[MS(P)-MS(PI)-MS(PR)+MS(PIR)]/ n_i n_R$ | $\sigma^2_{PIR} + n_i \sigma^2_{PR} + n_R \sigma^2_{PI} + n_i n_R \sigma^2_P$ |
| I | $[MS(I)-MS(PI)-MS(IR)+MS(PIR)]/ n_p n_R$ | $\sigma^2_{PIR} + n_p \sigma^2_{IR} + n_R \sigma^2_{PI} + n_p n_R \sigma^2_I$ |
| R | $[MS(R)-MS(PR)-MS(ER)+MS(PIR)]/ n_p n_i$ | $\sigma^2_{PIR} + n_p \sigma^2_{IR} + n_i \sigma^2_{PR} + n_p n_i \sigma^2_P$ |
| PI | $[MS(PI)-MS(PIR)]/ n_R$ | $\sigma^2_{PIR} + n_R \sigma^2_{PI}$ |
| PR | $[MS(PR)-MS(PIR)]/ n_i$ | $\sigma^2_{PIR} + n_i \sigma^2_{PR}$ |
| IR | $[MS(IR)-MS(PIR)]/ n_p$ | $\sigma^2_{PIR} + n_p \sigma^2_{IR}$ |
| PIR | $[MS(PIR)]$ | σ^2_{PIR} |

เมื่อ P แทน นักเรียน
I แทน งานเขียน
R แทน ผู้ประเมิน
PIR แทน ส่วนที่เหลือ

PI แทน นักเรียนกับงานเขียน
PR แทน นักเรียนกับผู้ประเมิน
IR แทน งานเขียนกับผู้ประเมิน

D-Study เพื่อหาค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) โดยใช้สูตรดังนี้

$$\sigma^2_{Rel} = (\sigma^2_{PR}/n_R) + (\sigma^2_{PI}/n_R) + (\sigma^2_{PIR}/n_R n_i)$$

$$\sigma^2_{Abs} = (\sigma^2_R/n_R) + (\sigma^2_I/n_i) + (\sigma^2_{PR}/n_R) + (\sigma^2_{PI}/n_i) + (\sigma^2_{IR}/n_R n_i) + (\sigma^2_{PIR}/n_R n_i)$$

$$\rho^2_{Rel} = (\sigma^2_P / \sigma^2_P + \sigma^2_{Rel})$$

$$\rho^2_{Abs} = (\sigma^2_P / \sigma^2_P + \sigma^2_{Abs})$$

เมื่อ P แทน นักเรียน σ^2_{Rel} แทน Relative Error Variance
I แทน งานเขียน σ^2_{Abs} แทน Absolute Error Variance
R แทน ผู้ประเมิน ρ^2_{Rel} แทน Relative G-Coefficient
 ρ^2_{Abs} แทน Absolute G-Coefficient

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรูปอ้างอิงของคะแนนการประเมินงานเขียนจากผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) และเพื่อหาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)

ในบทนี้ผู้วิจัยได้นำเสนอผลการวิเคราะห์ข้อมูลที่ได้จากการดำเนินการวิจัยตามลำดับดังนี้

1. ค่าสถิติพื้นฐานของคะแนนที่ได้จากการประเมินงานเขียนจากผู้ประเมินจำนวน 8 คน
2. การหาค่าความเที่ยงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนนจากค่าสัมประสิทธิ์การสรูปอ้างอิง (Generalizability Coefficient) มีขั้นตอนดังนี้
 - 2.1 การประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรูปอ้างอิงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนน
 - 2.2 คะแนนการศึกษาเพื่อการตัดสินใจการสรูปอ้างอิง เลือกรายงานผู้ประเมินและจำนวนงานเขียนที่เหมาะสมกับเกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมินจำแนกตามผู้ประเมินและตามรูปแบบเกณฑ์การให้คะแนน

ผลการวิเคราะห์ข้อมูลตามตาราง 2 พบว่า ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมินแต่ละคนเมื่อใช้รูปแบบเกณฑ์การให้คะแนนแบบภาพรวมมีดังนี้ ค่าเฉลี่ยเท่ากับ 11.1282, 7.3077, 9.9487 และ 10.1282 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.8639, 1.2598, 0.8568 และ 1.1045 ตามลำดับ และค่าเฉลี่ยรวมมีค่าเท่ากับ 9.6282 ส่วนเบี่ยงเบนมาตรฐานรวมมีค่าเท่ากับ 1.0213

และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าเฉลี่ยเป็น 9.2051, 8.6154, 8.9487 และ 9.5641 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 1.7944, 1.7713, 1.1909 และ 1.5355 ตามลำดับ และคะแนนเฉลี่ยรวมมีค่าเท่ากับ 9.0833 ส่วนเบี่ยงเบนมาตรฐานรวมมีค่าเท่ากับ 1.5730

ตาราง 2 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการให้คะแนนของผู้ประเมิน จำแนกตามผู้ประเมิน และตามรูปแบบเกณฑ์การให้คะแนน

| รูปแบบเกณฑ์การให้คะแนน | ผู้ประเมินคนที่ | คะแนนเต็ม | \bar{X} | S.D. |
|--|-----------------|-----------|-----------|--------|
| ประเมินรวม (holistic scoring rubric) | 1 | 12 | 11.1282 | 0.8639 |
| | 2 | 12 | 7.3077 | 1.2598 |
| | 3 | 12 | 9.9487 | 0.8568 |
| | 4 | 12 | 10.1282 | 1.1045 |
| | รวม | 12 | 9.6282 | 1.0213 |
| แยกองค์ประกอบ (analytic scoring rubric) | 1 | 12 | 9.2051 | 1.7944 |
| | 2 | 12 | 8.6154 | 1.7713 |
| | 3 | 12 | 8.9487 | 1.1909 |
| | 4 | 12 | 9.5641 | 1.5355 |
| | รวม | 12 | 9.0833 | 1.5730 |

2. การหาค่าความเที่ยงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนน จากค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) มีขั้นตอนดังนี้

2.1 ผลการประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability coefficient) ของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนน

ผู้วิจัยวิเคราะห์ข้อมูลมีจุดมุ่งหมายเพื่อคำนวณหาความแปรปรวนตามแหล่งความแปรปรวนต่างๆในเอกภพค่าสังเกตที่เป็นไปได้ทั้งหมดที่แยกเป็นอิสระจากกันขององค์ประกอบ 3

องค์ประกอบ ได้แก่ นักเรียน (P) งานเขียน (I) และผู้ประเมิน (R) รวมทั้งผลร่วมระหว่างนักเรียนกับงานเขียน (P) ผลร่วมระหว่างนักเรียนกับผู้ประเมิน (PR) ผลร่วมระหว่างงานเขียนกับผู้ประเมิน (IR) และส่วนที่เหลือ (PIR) ของคะแนนการประเมินงานเขียน เมื่อใช้รูปแบบเกณฑ์การให้คะแนนในแต่ละรูปแบบ

ผลการวิเคราะห์ข้อมูลตามตาราง 3 พบว่า การประเมินงานเขียนเมื่อใช้รูปแบบเกณฑ์การให้คะแนนแบบภาพรวม ค่าความแปรปรวนของแหล่งต่างๆ เมื่อเทียบกับค่าความแปรปรวนรวมมีค่าเรียงลำดับจากมากไปน้อยดังนี้ ความแปรปรวนผลร่วมระหว่างงานเขียนกับผู้ประเมิน มีค่า 0.29440 คิดเป็นร้อยละ 50.4282 ซึ่งแสดงให้เห็นถึงความแตกต่างของการให้คะแนนของผู้ประเมินในการประเมินงานเขียนแต่ละชิ้น รองลงมาคือความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจนมีค่า 0.23622 คิดเป็นร้อยละ 40.4591 ความแปรปรวนของนักเรียนมีค่า 0.04742 คิดเป็นร้อยละ 8.1192 ซึ่งแสดงถึงความแตกต่างของคะแนนอันเนื่องมาจากความสามารถของนักเรียนแต่ละคน

ตาราง 3 ผลการประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) ($n_p = 39$, $n_i = 3$, $n_r = 4$)

| Source of Variation | DF | Sum of Square (SS) | Mean Square (MS) | Estimated Variance Component | % of total variance |
|---------------------|-----|--------------------|------------------|------------------------------|---------------------|
| P | 38 | 4845.5833 | 0.8284 | 0.04742 | 8.1192 |
| I | 2 | 4837.5577 | 11.7265 | 0.0000 | 0.0000 |
| R | 3 | 4826.8291 | 4.2415 | 0.0000 | 0.0000 |
| PI | 76 | 4888.7500 | 0.2594 | 0.00579 | 0.9935 |
| PR | 114 | 4883.6667 | 0.2225 | 0.0000 | 0.0000 |
| IR | 6 | 4920.5897 | 11.7180 | 0.29440 | 50.4282 |
| PIR | 228 | 5051.0000 | 0.2362 | 0.23622 | 40.4591 |
| TOTAL | 467 | 34253.97649 | | 0.5838 | 100.0000 |

เมื่อ P แทน นักเรียน PI แทน นักเรียนกับงานเขียน
 I แทน งานเขียน PR แทน นักเรียนกับผู้ประเมิน
 R แทน ผู้ประเมิน IR แทน งานเขียนกับผู้ประเมิน
 PIR แทน ส่วนที่เหลือ

ผลการวิเคราะห์ข้อมูลตามตาราง 4 พบว่า การประเมินงานเขียนเมื่อใช้รูปแบบเกณฑ์การให้คะแนนแบบแยกองค์ประกอบค่าความแปรปรวนของแหล่งต่างๆเมื่อเทียบกับค่าความแปรปรวนมีค่าเรียงลำดับจากมากไปหาน้อยดังนี้ ความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจนมีค่า 0.36160 คิดเป็นร้อยละ 71.2625 รองลงมาคือความแปรปรวนของนักเรียน มีค่า 0.13135 คิดเป็นร้อยละ 25.8859 ซึ่งแสดงให้เห็นถึงความแตกต่างของนักเรียนแต่ละคนและความแปรปรวนผลร่วมระหว่างงานเขียนและผู้ประเมินมีค่า 0.01091 คิดเป็นร้อยละ 2.1500 ซึ่งแสดงให้เห็นถึงความแตกต่างของงานเขียนที่ได้จากการให้คะแนนของผู้ประเมินแต่ละคน

ตาราง 4 ผลการประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) ($n_p=39$, $n_r=3$, $n_f=4$)

| Source of Variation | DF | Sum of Square (SS) | Mean Square (MS) | Estimated Variance Component | % of total variance |
|---------------------|-----|--------------------|------------------|------------------------------|---------------------|
| P | 38 | 4370.3333 | 1.94512 | 0.13135 | 25.8859 |
| I | 2 | 4298.5513 | 1.06624 | 0.00174 | 0.3429 |
| R | 3 | 4297.0427 | 0.20798 | 0.00000 | 0.0000 |
| PI | 76 | 4400.5000 | 0.36887 | 0.00182 | 0.3587 |
| PR | 114 | 4408.6667 | 0.33078 | 0.00000 | 0.0000 |
| IR | 6 | 4303.8974 | 0.78704 | 0.01091 | 2.1500 |
| PIR | 228 | 4526.0000 | 0.36160 | 0.36160 | 71.2625 |
| TOTAL | 467 | 30604.99 | | 0.50742 | 100.0000 |

เมื่อ P แทน นักเรียน PI แทน นักเรียนกับงานเขียน
 I แทน งานเขียน PR แทน นักเรียนกับผู้ประเมิน
 R แทน ผู้ประเมิน IR แทน งานเขียนกับผู้ประเมิน
 PIR แทน ส่วนที่เหลือ

2.2 การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

ผู้วิจัยหาค่าความเที่ยงของคะแนนการให้คะแนนงานเขียนจากค่าสัมประสิทธิ์การสรุปอ้างอิง โดยใช้โปรแกรม GENOVA VERSION 2.1 ที่ได้จากเกณฑ์การให้คะแนนทั้งสองรูปแบบ แผน PXIXR จำนวนนักเรียน 39 คน จำนวนงานเขียน 3 ชิ้น และจำนวนผู้ประเมิน 4 คน ($n_p=39$, $n_i=3$, $n_r=4$)

ผลการวิเคราะห์ข้อมูลตามตาราง 5 พบว่าค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมมีค่าเท่ากับ 0.71484 และค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบมีค่าเท่ากับ 0.81410

ตาราง 5 ผลการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

| รูปแบบเกณฑ์การให้คะแนน | เอกภพของการสรุปอ้างอิง | σ^2_P | σ^2_{Rel} | ρ^2 |
|------------------------|------------------------|--------------|------------------|----------|
| แบบภาพรวม | P, I, R Random | 0.04935 | 0.01969 | 0.71484 |
| แบบแยกองค์ประกอบ | P, I, R Random | 0.13196 | 0.03013 | 0.81410 |

2.3 การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-STUDY) เลือกจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

การวิเคราะห์ขั้นตอนนี้เป็นการใช้ข้อมูลจากการศึกษาเพื่อการสรุปอ้างอิงมาศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-STUDY) ผู้วิจัยต้องการลดความคลาดเคลื่อนและเพิ่มความแม่นยำในการสรุปอ้างอิงโดยมีเงื่อนไขในการวัดคือ จำนวนงานเขียนและจำนวนผู้ประเมิน ให้จำนวนงานเขียนอยู่ระหว่าง 4-28 ชิ้น และจำนวนผู้ประเมิน 1-6 คน ซึ่งเป็นจำนวนที่อยู่ในวิสัยที่สามารถนำไปปฏิบัติได้ โดยการตัดสินใจจากการวิเคราะห์ว่า จำนวนงานเขียนกี่ชิ้นและจำนวนผู้ประเมินเท่าใด

ให้ค่าสัมประสิทธิ์การสรุปอ้างอิงสูงสุด เพื่อให้เห็นการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงที่ได้จากการตรวจให้คะแนนทั้ง 2 รูปแบบอย่างชัดเจน ผู้วิจัยได้นำเสนอปรากฏผลดังตาราง

ผลการวิเคราะห์ข้อมูลตามตาราง 6 พบว่า ถ้ากำหนดผู้ประเมินหนึ่งคน จำนวนงานเขียนจำนวน 4, 8, 12, 16, 20, 24 และ 28 ชิ้น ตรวจให้คะแนนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.26122, 0.41423, 0.58580, 0.63871, 0.67964 และ 0.71223 ตามลำดับ และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.58283, 0.73644, 0.80737, 0.89342 และ 0.9723 ตามลำดับ

ตาราง 6 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินหนึ่งคน ($n_p=39$, $n_t=4-28$, $n_r=1$)

| เกณฑ์การให้คะแนน | ESTIMATED VARIANCE COMPONENTS IN D-STUDY | | | | | | | |
|------------------|--|---------|---------|---------|---------|---------|---------|---------|
| | n_t | 4 | 8 | 12 | 16 | 20 | 24 | 28 |
| | n_r | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| แบบภาพรวม | σ_p^2 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 |
| | σ_{Abs}^2 | 0.13411 | 0.06705 | 0.04470 | 0.03353 | 0.02682 | 0.02235 | 0.01916 |
| | ρ_{Abs}^2 | 0.26122 | 0.41423 | 0.51473 | 0.58580 | 0.63871 | 0.67964 | 0.71223 |
| แบบแยกองค์ประกอบ | σ_p^2 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 |
| | σ_{Abs}^2 | 0.09402 | 0.04701 | 0.03134 | 0.02350 | 0.01880 | 0.01567 | 0.01343 |
| | ρ_{Abs}^2 | 0.58283 | 0.73644 | 0.80737 | 0.84822 | 0.87478 | 0.89342 | 0.90723 |

ผลการวิเคราะห์ข้อมูลตามตาราง 7 พบว่า ถ้ากำหนดผู้ประเมินสองคน จำนวนงานเขียน 4, 8, 12, 16, 20, 24 และ 28 ชิ้น ตรวจให้คะแนนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.41163, 0.58319, 0.67729, 0.73673, 0.77768, 0.80760 และ 0.8304 ตามลำดับ และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.73461, 0.84700, 0.89252, 0.91716, 0.93262, 0.9432 และ 0.95092 ตามลำดับ

ตาราง 7 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินสองคน ($n_p=39$, $n_i=4-28$, $n_r=2$)

| | | ESTIMATED VARIANCE COMPONENTS IN D-STUDY | | | | | | |
|------------------|------------------|--|---------|---------|---------|---------|---------|---------|
| เกณฑ์การให้คะแนน | n_i | 4 | 8 | 12 | 16 | 20 | 24 | 28 |
| | n_r | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| แบบภาพรวม | σ_p^2 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 |
| | σ_{Abs}^2 | 0.06778 | 0.03389 | 0.02259 | 0.01694 | 0.01356 | 0.01130 | 0.00968 |
| | ρ_{Abs}^2 | 0.41163 | 0.58319 | 0.67729 | 0.73673 | 0.77768 | 0.80760 | 0.83043 |
| แบบแยกองค์ประกอบ | σ_p^2 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 |
| | σ_{Abs}^2 | 0.04745 | 0.02373 | 0.01582 | 0.01186 | 0.00949 | 0.00791 | 0.00678 |
| | ρ_{Abs}^2 | 0.73461 | 0.84700 | 0.89252 | 0.91716 | 0.93262 | 0.94321 | 0.95092 |

ผลการวิเคราะห์ข้อมูลตามตาราง 8 พบว่าถ้ากำหนดผู้ประเมินสามคน จำนวนงานเขียน 4, 8, 12, 16, 20, 24 และ 28 ชิ้น ตรวจให้คะแนนเมื่อใช้เกณฑ์การให้คะแนนภาพรวมประมาณ ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.50940, 0.67497, 0.75698, 0.80595, 0.83849, 0.86168 และ 0.87905 ตามลำดับ และเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบประมาณ ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.80444, 0.89162, 0.92504, 0.94271, 0.95363, 0.96106 และ 0.96644 ตามลำดับ

ตาราง 8 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และ ค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้เกณฑ์การให้คะแนนแบบ ภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีมีผู้ประเมินสามคน ($n_p=39$, $n_i=4-28$, $n_r=3$)

| เกณฑ์การ ให้คะแนน | ESTIMATED VARIANCE COMPONENTS IN D-STUDY | | | | | | | |
|----------------------|--|---------|---------|---------|---------|---------|---------|---------|
| | n_i | 4 | 8 | 12 | 16 | 20 | 24 | 28 |
| | n_r | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| แบบภาพรวม | σ_p^2 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 |
| | σ_{Abs}^2 | 0.04567 | 0.02283 | 0.01522 | 0.01142 | 0.00913 | 0.00761 | 0.00652 |
| | ρ_{Abs}^2 | 0.50940 | 0.67497 | 0.75698 | 0.80595 | 0.83849 | 0.86168 | 0.87905 |
| แบบแยก องค์ประกอบ | σ_p^2 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 |
| | σ_{Abs}^2 | 0.03193 | 0.01597 | 0.01064 | 0.00798 | 0.00639 | 0.00532 | 0.00456 |
| | ρ_{Abs}^2 | 0.80444 | 0.89162 | 0.92504 | 0.94271 | 0.95363 | 0.96106 | 0.96644 |

ผลการวิเคราะห์ข้อมูลตามตาราง 9 พบว่าถ้ากำหนดผู้ประเมินสี่คน จำนวนงานเขียน 4, 8, 12, 16, 20, 24 และ 28 ชิ้น ตรวจสอบให้คะแนนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.57805, 0.73261, 0.80430, 0.84567, 0.87261, 0.89154 และ 0.90557 ตามลำดับ เมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.84458, 0.91574, 0.94220, 0.95602, 0.96450, 0.97624 และ 0.97438 ตามลำดับ

ตาราง 9 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีผู้ประเมินสี่คน ($n_p=39$, $n_i=4-28$, $n_r=4$)

| เกณฑ์การให้คะแนน | ESTIMATED VARIANCE COMPONENTS IN D-STUDY | | | | | | | |
|------------------|--|---------|---------|---------|---------|---------|---------|---------|
| | n_i | 4 | 8 | 12 | 16 | 20 | 24 | 28 |
| แบบภาพรวม | σ_p^2 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 |
| | σ_{Abs}^2 | 0.03461 | 0.01731 | 0.01154 | 0.00865 | 0.00692 | 0.00577 | 0.00494 |
| | ρ_{Abs}^2 | 0.57805 | 0.73261 | 0.80430 | 0.84567 | 0.87261 | 0.89154 | 0.90557 |
| แบบแยกองค์ประกอบ | σ_p^2 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 |
| | σ_{Abs}^2 | 0.02417 | 0.01209 | 0.00806 | 0.00604 | 0.00483 | 0.00403 | 0.00345 |
| | ρ_{Abs}^2 | 0.84458 | 0.91574 | 0.94220 | 0.95602 | 0.96405 | 0.97024 | 0.97438 |

ผลการวิเคราะห์ข้อมูลตามตาราง 10 พบว่าถ้ากำหนดผู้ประเมินห้าคน จำนวนงานเขียน 4, 8, 12, 16, 20, 24 และ 28 ชิ้น ตรวจให้คะแนนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.62890, 0.77218, 0.83564, 0.8144, 0.89444, 0.91046 และ 0.92226 ตามลำดับ และ เมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ประเมินค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.87065, 0.93085, 0.95281, 0.96419, 0.97114 , 0.97584 และ 0.97922 ตามลำดับ

ตาราง 10 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และ ค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีผู้ประเมินห้าคน ($n_p=39$, $n_t=4-28$, $n_r=5$)

| | | ESTIMATED VARIANCE COMPONENTS IN D-STUDY | | | | | | | |
|----------------------|------------------|--|---------|---------|---------|---------|---------|---------|--|
| เกณฑ์การ ให้คะแนน | n_t | 4 | 8 | 12 | 16 | 20 | 24 | 28 | |
| | n_r | 5 | 5 | 5 | 5 | 5 | 5 | 5 | |
| แบบภาพรวม | σ_p^2 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | |
| | σ_{Abs}^2 | 0.02798 | 0.01399 | 0.00933 | 0.00699 | 0.00560 | 0.00466 | 0.00400 | |
| | ρ_{Abs}^2 | 0.62890 | 0.77218 | 0.83564 | 0.87144 | 0.89444 | 0.91046 | 0.92226 | |
| แบบแยก องค์ประกอบ | σ_p^2 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | |
| | σ_{Abs}^2 | 0.01952 | 0.00976 | 0.00651 | 0.00488 | 0.00390 | 0.00325 | 0.00279 | |
| | ρ_{Abs}^2 | 0.87065 | 0.93085 | 0.95281 | 0.96419 | 0.97114 | 0.97584 | 0.97922 | |

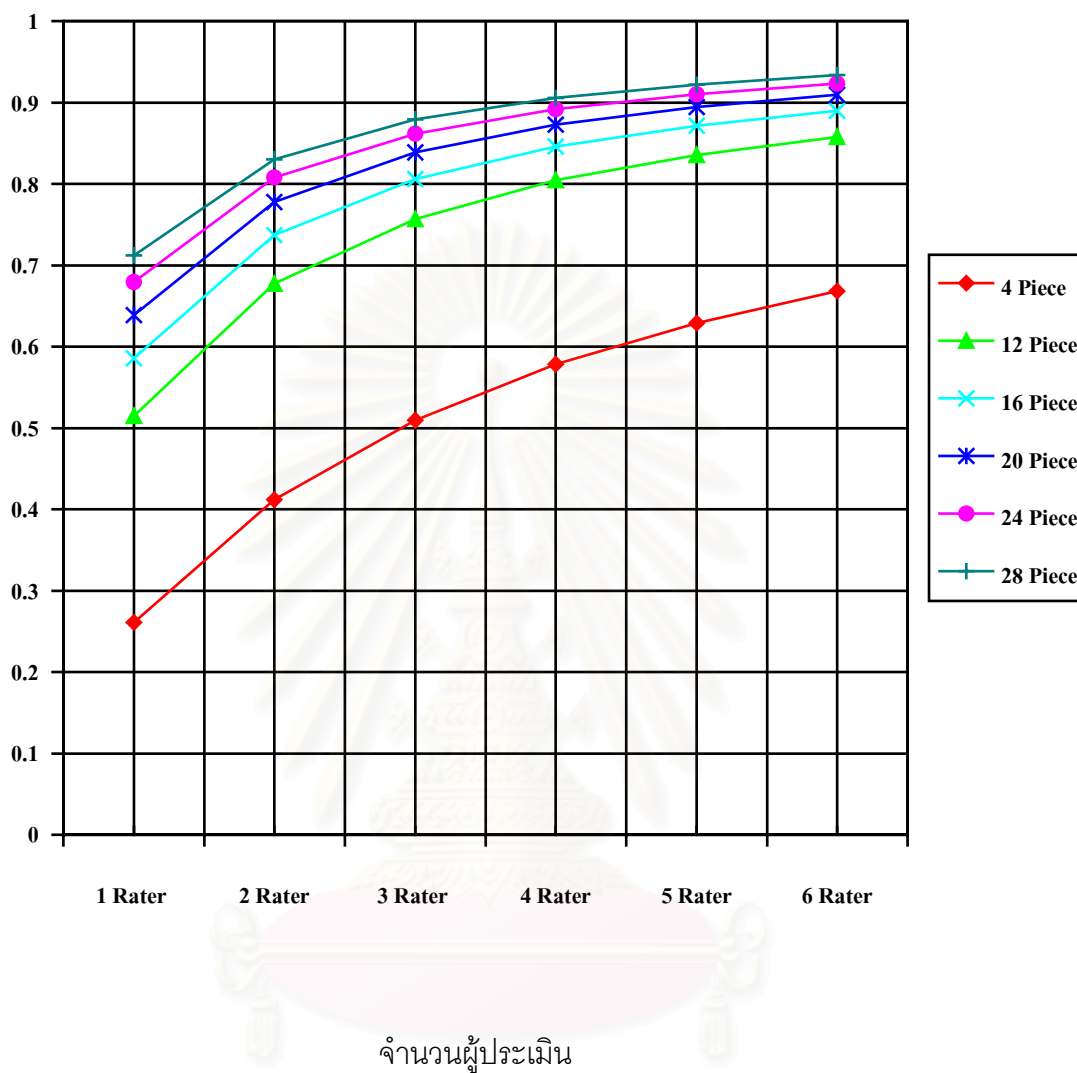
ผลการวิเคราะห์ข้อมูลตามตาราง 11 พบว่าถ้ากำหนดผู้ประเมินหกคน จำนวนงานเขียนจำนวน 4, 8, 12, 16, 20, 24 และ 28 ชิ้น ตรวจสอบให้คะแนนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.66808, 0.80102, 0.85792, 0.88952, 0.90962, 0.92353 และ 0.93373 ตามลำดับ และเมื่อเกณฑ์การให้คะแนนแบบแยกองค์ประกอบประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.88894, 0.94120, 0.96002, 0.96971, 0.97562, 0.97960 และ 0.98246 ตามลำดับ

ตาราง 11 แสดงการเปรียบเทียบค่าความแปรปรวนของคะแนนเอกภพ ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ กรณีผู้ประเมินหกคน ($n_p=39$, $n_i=4-28$, $n_r=6$)

| | | ESTIMATED VARIANCE COMPONENTS IN D-STUDY | | | | | | | |
|------------------|------------------|--|---------|---------|---------|---------|---------|---------|--|
| เกณฑ์การให้คะแนน | n_i | 4 | 8 | 12 | 16 | 20 | 24 | 28 | |
| | n_r | 6 | 6 | 6 | 6 | 6 | 6 | 6 | |
| แบบภาพรวม | σ_p^2 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | 0.04742 | |
| | σ_{Abs}^2 | 0.02356 | 0.01178 | 0.00785 | 0.00589 | 0.00471 | 0.00393 | 0.00337 | |
| | ρ_{Abs}^2 | 0.66808 | 0.80102 | 0.85792 | 0.88952 | 0.90962 | 0.92353 | 0.93373 | |
| แบบแยกองค์ประกอบ | σ_p^2 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | 0.13135 | |
| | σ_{Abs}^2 | 0.01641 | 0.00821 | 0.00547 | 0.00410 | 0.00328 | 0.00274 | 0.00234 | |
| | ρ_{Abs}^2 | 0.88894 | 0.94120 | 0.96002 | 0.96971 | 0.97562 | 0.97960 | 0.98246 | |

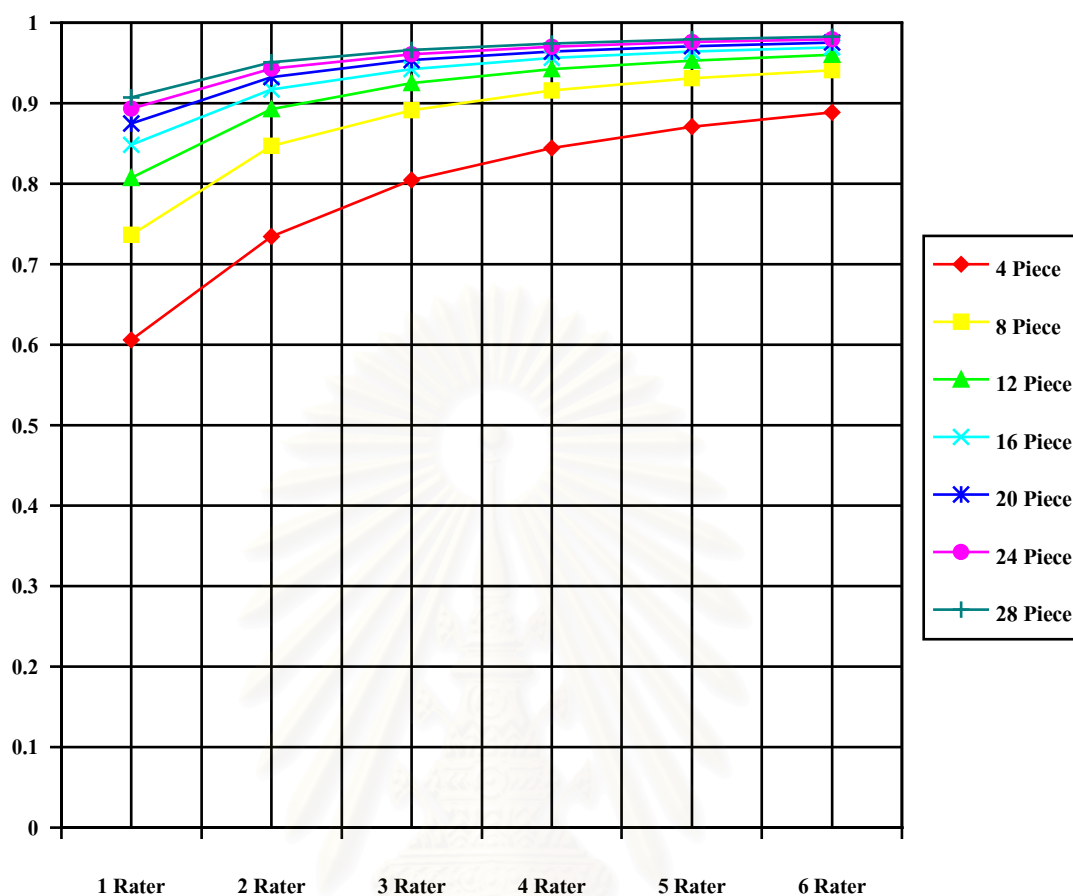
เพื่อแสดงให้เห็นการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการให้คะแนนงานเขียนที่ใช้เกณฑ์การให้คะแนนทั้งสองรูปแบบตามรายละเอียดที่กล่าวมาข้างต้นอย่างชัดเจน ผู้วิจัยนำเสนอการเปรียบเทียบเป็นกราฟเส้น ดังแสดงในภาพประกอบ

ค่าสัมประสิทธิ์การสรุปอ้างอิง



แผนภาพ 3 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนที่ได้จากการใช้เกณฑ์การให้คะแนนแบบภาพรวม ที่จำนวนงานเขียน 4-28 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามแบบแผนการวัด P X I R

ค่าสัมประสิทธิ์การสรุปอ้างอิง



จำนวนผู้ประเมิน

แผนภาพ 4 แสดงการเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนที่ได้จากการใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ที่จำนวนงานเขียน 4-28 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามแบบแผนการวัด P X I X R

จากตาราง 6-11 และแผนภาพ 3-4 พบว่าเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ช่วงการตัดสินใจที่จำนวนงานเขียน 4-28 ขึ้น จำนวนผู้ประเมินตั้งแต่ 1-6 คน ให้ค่าสัมประสิทธิ์การสรุปอ้างอิงสูงกว่า เกณฑ์การให้คะแนนแบบภาพรวม ทุกช่วงค่าโดยเกณฑ์การให้คะแนนแบบแยกองค์ประกอบมีค่าระหว่าง 0.58283 ถึง 0.98246 และเกณฑ์การให้คะแนนแบบภาพรวม มีค่าระหว่าง 0.26122 ถึง 0.93373 จากกราฟจะเห็นว่าเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ความชันของกราฟเริ่มเปลี่ยนแปลงน้อยลง เมื่อจำนวนชิ้นงานตั้งแต่ 12 ขึ้นขึ้นไป สำหรับเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ความชันของกราฟเริ่มเปลี่ยนแปลงน้อยลง เมื่อจำนวนชิ้นงานตั้งแต่ 8 ขึ้นขึ้นไป นั่นแสดงว่า ถึงแม้ว่าจำนวนชิ้นงานจะมากขึ้นก็ไม่มีผลต่อการเพิ่มค่าสัมประสิทธิ์การสรุปอ้างอิงมากนัก

ช่วงการศึกษาเพื่อการตัดสินใจ (D-STUDY) จำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสม เมื่อใช้ผู้ประเมิน 1 คน ซึ่งเป็นวิธีที่ปฏิบัติในการเรียนการสอนทั่วไปและกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2, 0.4, 0.6 และ 0.8 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ต้องใช้จำนวนชิ้นงานอย่างน้อย 4, 8, 16 และ 28 ตามลำดับ ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 และ 0.4 ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 ใช้จำนวนชิ้นงานอย่างน้อย 8 ชิ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.8 ใช้จำนวนชิ้นงานอย่างน้อย 12 ชิ้น

เมื่อใช้ผู้ประเมิน 2 คน หรือ 3 คน ซึ่งสามารถปฏิบัติได้และกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 และสำหรับเกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.4, 0.6 และ 0.8 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ใช้จำนวนชิ้นงานอย่างน้อย 4, 12 และ 24 ขึ้น ตามลำดับ เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.4 สำหรับเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 และ 0.8 สำหรับเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานอย่างน้อย 4 และ 8 ขึ้น ตามลำดับ

เมื่อใช้ผู้ประเมิน 4 คน และกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 สำหรับเกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.4 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ใช้จำนวนชิ้นงานอย่างน้อย 4 ขึ้น ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ขึ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ใช้จำนวนชิ้นงานอย่างน้อย 12 ขึ้น ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานอย่างน้อย 4 ขึ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น

0.8 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ใช้จำนวนชิ้นงานอย่างน้อย 16 ชิ้น ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานอย่างน้อย 8 ชิ้น

เมื่อใช้ผู้ประเมิน 5 หรือ 6 คน และกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 สำหรับเกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.4 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ใช้จำนวนชิ้นงานอย่างน้อย 4 ชิ้น ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 และ 0.8 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ใช้จำนวนชิ้นงานอย่างน้อย 12 และ 16 ชิ้น ตามลำดับ ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานอย่างน้อย 4 ชิ้น และ 8 ชิ้น ตามลำดับ



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สรุปผลการวิจัย อภิปรายและข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความเที่ยงซึ่งแทนด้วยค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนจากผู้ประเมินจำนวน 4 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) และผู้ประเมินจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric) และเพื่อหาจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสมเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)

กลุ่มตัวอย่าง

กลุ่มตัวอย่าง ในการวิจัยครั้งนี้มีกลุ่มตัวอย่างที่ศึกษาประกอบด้วย

กลุ่มตัวอย่าง ผลงานการเขียนวิชาภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 5 ภาคเรียนที่ 1 ปีการศึกษา 2544 จำนวน 1 ห้องเรียน ของโรงเรียนไผทอุดมศึกษา

กลุ่มตัวอย่างผู้ประเมิน เนื่องจากในงานวิจัยครั้งนี้ต้องใช้ครูผู้สอนวิชาภาษาไทย ชั้นประถมศึกษาปีที่ 5-6 เป็นผู้ประเมินผลงานการเขียนซึ่งต้องใช้เวลาและความสนใจ ตั้งใจในการประเมินผลงานการเขียน ผู้วิจัยทำการสุ่มครูผู้ประเมินโดยวิธีการแบบเจาะจง (purposive sampling) เพราะต้องหาผู้ประเมินที่เต็มใจและมีเวลาช่วยประเมินอย่างเต็มที่ มีประสบการณ์ในการสอนอย่างน้อย 5 ปี ครูที่ได้รับเลือกเป็นกลุ่มตัวอย่างเป็นครูจาก โรงเรียนไผทอุดมศึกษา จำนวน 3 คน โรงเรียนอัสสัมชัญแผนกประถม จำนวน 3 คน และ โรงเรียนศิริวิทยา จำนวน 2 คน รวมจำนวนทั้งหมด 8 คน แบ่งครูออกเป็น 2 กลุ่มๆละ 4 คนโดยใช้วิธีสุ่มอย่างง่ายโดยการจับสลาก ซึ่งผู้ประเมินกลุ่มที่ 1 จะใช้เกณฑ์การให้คะแนนแบบภาพรวม และ ผู้ประเมินกลุ่มที่ 2 จะใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

เครื่องมือที่ใช้ในการศึกษา

เครื่องมือที่ใช้ในการศึกษาประกอบด้วย

- 1) เกณฑ์การให้คะแนนแบบภาพรวม
- 2) เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ
- 3) แบบบันทึกคะแนน

วิเคราะห์ข้อมูล

ผู้วิจัยนำข้อมูลที่ได้มาวิเคราะห์ด้วยโปรแกรม GENOVA เพื่อหาค่าความเที่ยงจากค่าสัมประสิทธิ์การสรุปอ้างอิง

สรุปผลการวิจัย

ผู้วิจัยได้ดำเนินการวิเคราะห์ข้อมูลออกเป็น 2 ตอน ดังนี้

ตอนที่ 1 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการประเมินงานเขียนจากผู้ประเมินจำนวน 8 คน

ตอนที่ 2 การหาค่าความเที่ยงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนนจากค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) มีขั้นตอนดังนี้

2.1 การประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนน

2.2 การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิงเลือกจำนวนผู้ประเมินและจำนวนงานเขียน ที่เหมาะสมกับเกณฑ์การให้คะแนนแบบภาพรวม (holistic scoring rubric) หรือเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ (analytic scoring rubric)

ตอนที่ 1 ค่าสถิติพื้นฐานของคะแนนที่ได้จากการประเมินงานเขียนจากผู้ประเมินจำนวน 8 คน

ค่าสถิติพื้นฐานของคะแนนที่ได้จากการประเมินงานเขียนของผู้ประเมินทั้งสี่คนเมื่อใช้รูปแบบเกณฑ์การให้คะแนนแบบภาพรวม มีดังนี้ ค่าเฉลี่ยเท่ากับ 11.1282, 7.3077, 9.9487 และ 10.1282 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 0.8639, 0.8568 และ 1.1045 ตามลำดับ และค่าเฉลี่ยร่วมมีค่าเท่ากับ 9.6282 ส่วนเบี่ยงเบนมาตรฐานร่วมมีค่าเท่ากับ 1.0213 และจากผู้ประเมินจำนวนสี่คนเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบมีค่าเฉลี่ยเป็น 9.2051, 8.6154, 8.9487 และ 9.5641 ตามลำดับ ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 1.7944, 1.7713, 1.1909 และ 1.5355 ตามลำดับ และคะแนนเฉลี่ยร่วมมีค่าเท่ากับ 9.0833 ส่วนเบี่ยงเบนมาตรฐานร่วมมีค่าเท่ากับ 1.5730

ตอนที่ 2 การหาค่าความเที่ยงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนนจากค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient)

1. การประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนน

1.1 แหล่งความแปรปรวนที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ได้แก่ ความแปรปรวนผลร่วมระหว่างงานเขียนกับผู้ประเมิน σ^2_{IR} มีค่ามากที่สุดคือ 0.29440 คิดเป็นร้อยละ 50.4282 รองลงมาคือความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจน σ^2_{PIR} มีค่าเท่ากับ 0.23622 คิดเป็นร้อยละ 40.4591 ความแปรปรวนของนักเรียน σ^2_p มีค่า 0.04742 คิดเป็นร้อยละ 8.1192 ความแปรปรวนผลร่วมระหว่างนักเรียนและงานเขียน σ^2_{PI} มีค่าเท่ากับ 0.00579 คิดเป็นร้อยละ 0.9935 ส่วนความแปรปรวนของงานเขียน σ^2_I ความแปรปรวนของผู้ประเมิน σ^2_R และความแปรปรวนผลร่วมระหว่างนักเรียนและผู้ประเมิน σ^2_{PR} มีค่าเป็น 0

1.2 แหล่งความแปรปรวนที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ได้แก่ ความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจน σ^2_{PIR} มีค่ามากที่สุดคือ 0.36160 คิดเป็นร้อยละ 71.2625 รองลงมาคือความแปรปรวนของนักเรียน σ^2_p มีค่า 0.13135 คิดเป็นร้อยละ 25.8859 ความแปรปรวนผลร่วมระหว่างงานเขียนกับผู้ประเมิน σ^2_{IR} มีค่าเท่ากับ 0.01091 คิดเป็นร้อยละ 2.1500 ส่วนความแปรปรวนของผู้ประเมิน σ^2_R และความแปรปรวนผลร่วมระหว่างนักเรียนและผู้ประเมิน σ^2_{PR} มีค่าเป็น 0

2. การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

คะแนนการประเมินงานเขียนเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมมีค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ 0.71484 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบมีค่าสัมประสิทธิ์การสรุปอ้างอิง เท่ากับ 0.81410

3. การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-STUDY) เลือกจำนวนผู้ประเมินและจำนวนงานเขียนที่เหมาะสม

เมื่อนำผลงานการเขียนจำนวน 117 ผลงาน จำนวนผู้ประเมิน 4 คน ประเมินผลงานการเขียนโดยใช้เกณฑ์การให้คะแนนแบบภาพรวมและผู้ประเมินอีกจำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบรวมผู้ประเมินจำนวนทั้งหมด 8 คน การศึกษาข้อมูลและวิเคราะห์ตามหลักของทฤษฎีการสรุปอ้างอิงที่จำนวนงานเขียนและจำนวนผู้ประเมิน ต่างกันให้ผลดังนี้

3.1 ค่าสัมประสิทธิ์การสรุปอ้างอิง ที่จำนวนงานเขียน 4 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามเกณฑ์การให้คะแนนแบบภาพรวม มีค่าตั้งแต่ 0.26122 ถึง 0.66808 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าตั้งแต่ 0.58283 ถึง 0.98246 จะเห็นได้ว่าเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม

3.2 ค่าสัมประสิทธิ์การสรุปอ้างอิง ที่จำนวนงานเขียน 8 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามเกณฑ์การให้คะแนนแบบภาพรวม มีค่าตั้งแต่ 0.41423 ถึง 0.80102 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าตั้งแต่ 0.73644 ถึง 0.94120 จะเห็นได้ว่าเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม

3.3 ค่าสัมประสิทธิ์การสรุปอ้างอิง ที่จำนวนงานเขียน 12 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามเกณฑ์การให้คะแนนแบบภาพรวม มีค่าตั้งแต่ 0.51473 ถึง 0.85792 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าตั้งแต่ 0.80737 ถึง 0.96002 จะเห็นได้ว่าเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม

3.4 ค่าสัมประสิทธิ์การสรุปอ้างอิง ที่จำนวนงานเขียน 16 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามเกณฑ์การให้คะแนนแบบภาพรวม มีค่าตั้งแต่ 0.58580 ถึง 0.88952 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าตั้งแต่ 0.84822 ถึง 0.96971 จะเห็นได้ว่าเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม

3.5 ค่าสัมประสิทธิ์การสรุปอ้างอิง ที่จำนวนงานเขียน 20 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามเกณฑ์การให้คะแนนแบบภาพรวม มีค่าตั้งแต่ 0.63871 ถึง 0.90962 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าตั้งแต่ 0.87478 ถึง 0.97562 จะเห็นได้ว่าเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม

3.6 ค่าสัมประสิทธิ์การสรุปอ้างอิง ที่จำนวนงานเขียน 24 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามเกณฑ์การให้คะแนนแบบภาพรวม มีค่าตั้งแต่ 0.67964 ถึง 0.92353 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าตั้งแต่ 0.89342 ถึง 0.97960 จะเห็นได้ว่าเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม

3.7 ค่าสัมประสิทธิ์การสรุปอ้างอิง ที่จำนวนงานเขียน 28 ชิ้น จำนวนผู้ประเมิน 1-6 คน ตามเกณฑ์การให้คะแนนแบบภาพรวม มีค่าตั้งแต่ 0.71223 ถึง 0.93373 และจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ มีค่าตั้งแต่ 0.90723 ถึง 0.98246 จะเห็นได้ว่าเกณฑ์การให้คะแนนแบบแยกองค์ประกอบให้ค่าสัมประสิทธิ์การสรุปอ้างอิงที่มีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม

อภิปรายผลการวิจัย

1. สำหรับคะแนนการประเมินงานเขียนจากจำนวนผู้ประเมิน 8 คน โดยผู้ประเมินในกลุ่มที่ 1 จำนวน 4 คน ที่ใช้เกณฑ์การให้คะแนนแบบภาพรวมในการประเมินผลงานและกลุ่มที่ 2 จำนวน 4 คนที่ใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ จากคะแนนผลการตรวจที่ผู้ประเมินใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบเมื่อผู้ประเมินคนที่ 1 มีคะแนนเฉลี่ยมีค่าเท่ากับ 9.2051 คิดเป็นร้อยละ 57.53 ผู้ประเมินคนที่ 2 มีคะแนนเฉลี่ยมีค่าเท่ากับ 8.6154 คิดเป็นร้อยละ 53.85 ผู้ประเมินคนที่ 3 มีคะแนนเฉลี่ยมีค่าเท่ากับ 8.9487 คิดเป็นร้อยละ 55.93 ผู้ประเมินคนที่ 4 มีคะแนนเฉลี่ยมีค่าเท่ากับ 9.5641 คิดเป็นร้อยละ 59.78 ซึ่งมีค่าใกล้เคียงกัน

ส่วนผู้ประเมินที่ใช้เกณฑ์การให้คะแนนภาพรวมเมื่อผู้ประเมินคนที่ 1 มีคะแนนเฉลี่ยมีค่าเท่ากับ 11.1282 คิดเป็นร้อยละ 69.55 ผู้ประเมินคนที่ 2 มีคะแนนเฉลี่ยมีค่าเท่ากับ 7.3077 คิดเป็นร้อยละ 45.67 ผู้ประเมินคนที่ 3 มีคะแนนเฉลี่ยมีค่าเท่ากับ 9.9487 คิดเป็นร้อยละ 62.18 ผู้ประเมินคนที่ 4 มีคะแนนเฉลี่ยมีค่าเท่ากับ 10.1282 คิดเป็นร้อยละ 63.30 มีค่าแตกต่างกัน

2. การประมาณค่าความแปรปรวนขององค์ประกอบ 3 องค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนในแต่ละรูปแบบเกณฑ์การให้คะแนน จากการศึกษาอิทธิพลขององค์ประกอบต่างๆที่เกี่ยวข้องกับคะแนนการประเมินงานเขียนที่ได้จากเกณฑ์การให้คะแนนทั้งสองรูปแบบ การวิจัยครั้งนี้ผู้วิจัยได้ศึกษาองค์ประกอบ 3 องค์ประกอบ คือ นักเรียน (P) งานเขียน (I) และผู้ประเมิน (R) ซึ่งกำหนดให้ความสัมพันธ์เป็นแบบ PXIXR ในเกณฑ์การให้คะแนนแบบภาพรวม แหล่งความแปรปรวนผลร่วมระหว่างงานเขียนกับ

ผู้ประเมิน เป็นแหล่งความแปรปรวนที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงมากที่สุด แสดงให้เห็นว่าทักษะการใช้เกณฑ์การให้คะแนนของผู้ประเมินร่วมกับจำนวนงานเขียนในแต่ละชั้น ส่งผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง และพบว่าผู้ประเมินแต่ละคนตรวจให้คะแนนงานเขียนแต่ละชิ้นงานสอดคล้องกัน สังเกตได้จากค่าความแปรปรวนของผู้ประเมินมีค่าเป็นศูนย์ จึงเป็นการขจัดความแปรปรวนของผู้ประเมินได้อย่างสิ้นเชิง เนื่องจากผู้ประเมินกลุ่มนี้มีความเข้าใจตรงกันกับคำชี้แจงของการใช้เกณฑ์การให้คะแนนที่กำหนดไว้ทำให้ผลการตรวจของผู้ประเมินแต่ละคน สอดคล้องกันมากขึ้น ซึ่งสอดคล้องกับงานวิจัยของ Suzanne et al. (1997) ที่ว่าความคลาดเคลื่อนที่เนื่องมาจากผู้ประเมินจะมีค่าน้อยมาก เมื่อมีเกณฑ์การให้คะแนนที่ชัดเจน

ส่วนในเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ แหล่งความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจนเป็นแหล่งความแปรปรวนที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงมากที่สุด และค่าความแปรปรวนของผู้ประเมินมีค่าเป็นศูนย์

3. การประมาณค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินสมงานเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวมและเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

จากการศึกษาค่าสัมประสิทธิ์การสรุปอ้างอิงของคะแนนการประเมินงานเขียนค่าสัมประสิทธิ์การสรุปอ้างอิงเมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม มีค่าเท่ากับ 0.71484 และค่าสัมประสิทธิ์การสรุปอ้างอิงเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบมีค่าเท่ากับ 0.81410 แสดงว่าค่าความเที่ยงเมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบมีค่าสูงกว่าเกณฑ์การให้คะแนนแบบภาพรวม อาจเนื่องมาจากเกณฑ์การให้คะแนนแบบแยกองค์ประกอบในแต่ละประเด็นย่อยๆของแต่ละชิ้นงาน เกณฑ์การให้คะแนนมีลักษณะไม่ขึ้นต่อกัน ทำให้มีความเป็นปรนัย จากคำกล่าวของ Hopskin และ Stanley (1981 อ้างถึงใน อรุณี ไร่อรุณ, 2536) ที่ว่าการตรวจให้คะแนนโดยใช้วิธีการตรวจแบบแยกองค์ประกอบ จะให้ค่าความเที่ยงสูงกว่าวิธีการตรวจแบบภาพรวม ทั้งนี้เนื่องจากวิธีการตรวจแบบแยกองค์ประกอบจะมีระบบการให้คะแนนที่ชัดเจน มีการแยกเป็นประเด็นย่อยๆ ตามองค์ประกอบ และยังกำหนดน้ำหนักคะแนนในแต่ละส่วน ซึ่งข้อดีคือ คะแนนที่ได้นำเชื่อถือ การแบ่งคะแนนทำให้ง่ายต่อการตรวจ ทำให้ผู้ประเมินสามารถใช้เกณฑ์การให้คะแนนได้อย่างค่อนข้างเป็นปรนัย ในขณะที่วิธีการตรวจแบบภาพรวมถึงแม้จะมีเกณฑ์การให้คะแนนแต่ละระดับเช่นกัน แต่คำตอบไม่ได้ถูกแบ่งเป็นประเด็นย่อยๆ การให้คะแนนจึงขึ้นอยู่กับประเด็นหลัก โดยจะให้อิสระแก่ผู้ประเมินในการพิจารณาคำตอบ จึงทำให้เกิดความคลาดเคลื่อนของคะแนนได้มาก ซึ่งเมื่อคะแนนมีความคลาดเคลื่อนมาก จะส่งผลให้ค่าความเที่ยงที่ได้มีค่าน้อย

4. การศึกษาเพื่อการตัดสินใจการสรุปอ้างอิง (D-STUDY) เลือกจำนวนผู้ประเมินที่เหมาะสมกับรูปแบบเกณฑ์การให้คะแนน และจำนวนผลงานของการประเมินเพิ่มสะสมงาน

จากการศึกษาเพื่อการตัดสินใจ (D-STUDY) จำนวนผลงาน และจำนวนผู้ประเมินที่เหมาะสมสำหรับเกณฑ์การให้คะแนนแบบภาพรวม ถ้าต้องการใช้ผู้ประเมิน 1 คน ซึ่งเป็นวิธีที่ปฏิบัติในการเรียนการสอนทั่วไปและกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2, 0.4, 0.6 และ 0.8 สำหรับเกณฑ์การให้คะแนนแบบภาพรวม ต้องใช้จำนวนชิ้นงานอย่างน้อย 4, 8, 16 และ 28 ตามลำดับ ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 และ 0.4 ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 ใช้จำนวนชิ้นงานอย่างน้อย 8 ชิ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.8 ใช้จำนวนชิ้นงานอย่างน้อย 12 ชิ้น หรือ ถ้าต้องการใช้ผู้ประเมิน 2 คน หรือ 3 คน ซึ่งสามารถปฏิบัติได้ และกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.2 และสำหรับเกณฑ์การให้คะแนนแบบภาพรวม และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.4, 0.6 และ 0.8 สำหรับเกณฑ์การให้คะแนนแบบภาพรวมใช้จำนวนชิ้นงานอย่างน้อย 4, 12 และ 24 ชิ้น ตามลำดับ เมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.4 สำหรับเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานน้อยกว่า 4 ชิ้น และเมื่อกำหนดค่าสัมประสิทธิ์การสรุปอ้างอิงเป็น 0.6 และ 0.8 สำหรับเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ใช้จำนวนชิ้นงานอย่างน้อย 4 และ 8 ชิ้น ตามลำดับ ทั้งนี้ขึ้นอยู่กับ การตัดสินใจช่วยให้ผู้ปฏิบัติสามารถเลือกช่วงการตัดสินใจใช้เกณฑ์การให้คะแนนรูปแบบใด ที่นำไปปฏิบัติจริงแล้วได้ค่าความเที่ยงตามที่ต้องการ เช่น ถ้าผู้สอน 1 คน ต้องการค่าความเที่ยง 0.8 ในเกณฑ์การให้คะแนนแบบภาพรวม ต้องใช้ชิ้นงานอย่างน้อย 28 ชิ้น จะเห็นว่ามีจำนวนชิ้นงานที่สูงมากในความเป็นจริงอาจจะทำได้ แต่ไม่สะดวกและต้องใช้เวลานาน ดังนั้นผู้สอนอาจปรับค่าความเที่ยงเป็น 0.6 แทน ซึ่งใช้ผลงานอย่างน้อย 16 ชิ้น ทำให้สะดวกและประหยัดเวลาได้ โดยสรุปการพิจารณาว่าจะเพิ่มจำนวนชิ้นงานหรือจำนวนผู้ประเมินนั้น จะต้องพิจารณาว่าสามารถปฏิบัติได้จริง มีความสะดวกและประหยัด ในการเรียนการสอนเราสามารถเพิ่มจำนวนชิ้นงานได้มากกว่าเพิ่มจำนวนผู้ประเมิน โดยเฉพาะการใช้ครูจำนวนมาก เป็นสิ่งที่ทำได้ยากสำหรับการประเมินในโรงเรียน แต่การเพิ่มชิ้นงานสามารถทำได้ง่ายกว่า

ข้อเสนอแนะ

ข้อเสนอแนะในการนำไปใช้

1. ในการประเมินผลงานการเขียน ผู้ประเมินควรสร้างคู่มือการตรวจให้คะแนน และกำหนดเกณฑ์การให้คะแนนที่ชัดเจน
2. ครูผู้สอนสามารถนำเกณฑ์การให้คะแนนไปตรวจให้คะแนนผลงานการเขียน และสามารถเป็นข้อมูลย้อนกลับแก่นักเรียนว่ามีข้อบกพร่องอย่างไร
3. ในการประยุกต์ใช้ถ้าผู้ปฏิบัติต้องการยึดที่จำนวนชิ้นงาน ชิ้นงานนั้นจะต้องสามารถปฏิบัติได้ ใช้เวลาน้อย และเหมาะสมกับการเรียนการสอน ซึ่งใช้ผู้ประเมินจำนวนน้อย แต่ถ้าต้องการยึดจำนวนผู้ประเมิน ชิ้นงานควรมีจำนวนน้อย แต่ต้องใช้ผู้ประเมินจำนวนมาก เช่น ในการประเมินผลงานการเขียนของนักเรียนระดับชั้นประถมศึกษา ถ้าต้องการความเที่ยงสูง โดยใช้ผู้ประเมินจำนวน 1 คน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ควรใช้ชิ้นงานอย่างน้อย 8 ชิ้น ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ควรใช้ชิ้นงานอย่างน้อย 4 ชิ้น แต่ถ้าต้องการใช้ชิ้นงานจำนวน 1 ชิ้น เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม ควรใช้ผู้ประเมินอย่างน้อย 4 คน ส่วนเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ควรใช้ผู้ประเมินอย่างน้อย 2 คน

ข้อเสนอแนะในวิจัยครั้งต่อไป

1. ควรศึกษาการประเมินผลงาน โดยให้ผู้เรียนเป็นผู้สร้างเกณฑ์การให้คะแนน และประเมินผลงานนั้นด้วยเกณฑ์ที่สร้างขึ้น เพราะจากงานวิจัยครั้งนี้ใช้ครูเป็นผู้ประเมิน แต่การศึกษาในปัจจุบันเน้นการประเมินตามสภาพจริง จึงต้องให้ผู้เรียนประเมินผลงานด้วยตนเอง
2. ควรศึกษาองค์ประกอบอื่นๆที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง เช่น จำนวนครั้งในการตรวจให้คะแนน เพราะจะได้ตรวจสอบ ยืนยัน ข้อมูลที่ได้จากงานวิจัยครั้งนี้

รายการอ้างอิง

ภาษาไทย

- กนกวรรณ เอี่ยมชัย. (2539). การศึกษาค่าความเชื่อมั่นของแบบทดสอบเอ็มอีคิว วิชาแนวคิดพื้นฐานและหลักการพยาบาลที่ตรวจให้คะแนนต่างกัน. วิทยานิพนธ์การศึกษามหาบัณฑิต ภาควิชาวัดผลและวิจัยการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- จิราพร อัครสมพงศ์. (2536). การเปรียบเทียบความสอดคล้องของคะแนนแบบสอบความเรียงจากผู้ตรวจและวิธีการตรวจต่างกัน. วิทยานิพนธ์ปริญญาามหาบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
- ชัยพฤกษ์ เสรีรักษ์ และคณะ. (2540). คู่มือการประเมินผลโดยใช้แฟ้มผลงานดีเด่น โครงการพัฒนาทรัพยากรมนุษย์ จ.กระบี่. กรุงเทพมหานคร: สถาบันพัฒนาคุณภาพวิชาการ (พว.).
- ชัยฤทธิ์ ศิลาเดช. (2540). การพัฒนาแฟ้มสะสมงานในการประเมินผลการเรียนวิชาภาษาอังกฤษระดับมัธยมศึกษาปีที่ 3. วิทยานิพนธ์การศึกษาดุษฎีบัณฑิต ภาควิชาวัดผลและวิจัยการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- พรรณณี เจียมสุขบุตร. (2543). การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจและวิธีการตรวจต่างกัน. วิทยานิพนธ์การศึกษามหาบัณฑิต ภาควิชาวัดผลและวิจัยการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- ไพรัตน์ วงษ์นาม. (2533). สัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบความเรียง. วิทยานิพนธ์ปริญญาามหาบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
- วิชาการ,กรม. (2538). การประเมินผลตามสภาพจริง. กรุงเทพฯ: อัดสำเนา.
- วิชาการ,กรม. (2540). กระบวนการพัฒนาแฟ้มสะสมงานของนักเรียน. กรุงเทพฯ: อัดสำเนา.
- ภาวิณี ศรีสุขวัฒนานนท์. (2540). "การประเมินผลสัมฤทธิ์ด้วยระบบแฟ้มสะสมงาน". วารสารศึกษาศาสตร์. 12 (มกราคม-เมษายน):1-17.
- วาสนา ประवालพุกษ์. (2537). "การวัดผลจากการปฏิบัติจริง". วารสารวัดผลทางการศึกษา. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร 16 (มกราคม-เมษายน): 36-43.

- ศิริชัย กาญจนวาสี. (2542). **ทฤษฎีการวัดและประเมินผล**. เอกสารประกอบการสอน.
กรุงเทพมหานคร: ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
ศึกษาศึกษา, กระทรวง. (2542). **พระราชบัญญัติการศึกษาแห่งชาติ ฉบับเป็นกฎหมาย**.
กรุงเทพมหานคร: โรงพิมพ์คุรุสภาลาดพร้าว.
- สมชาย มิ่งมิตร. (2539). **ผลของการประเมินจากพอร์ตโฟลิโอที่มีต่อผลสัมฤทธิ์ทางการ
เรียนวิชาภาษาไทยของนักเรียนประถมศึกษาปีที่ 5**. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต
ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.
- สมศักดิ์ ภูวิภาดาธรรม. (2544). **การยึดผู้เรียนเป็นศูนย์กลางและการประเมินตามสภาพ
จริง**. เชียงใหม่: สำนักพิมพ์ The Knowledge Center.
- อมรรัตน์ พันธุ์งาม. (2543). **การพัฒนากระบวนการประเมินเพิ่มสะสมงาน และการใช้
ประโยชน์จากการประเมินเพิ่มสะสมงาน: การประยุกต์ใช้การประเมินอภิमान**.
วิทยานิพนธ์ปริญญาโทมหาบัณฑิต ภาควิชาวิจัยการศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์
มหาวิทยาลัย
- อุษณีย์ บัวศิริพันธุ์. (2543). **การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบ
ทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจ และประสบการณ์ของผู้
ผู้ตรวจ แตกต่างกัน**. วิทยานิพนธ์การศึกษามหาบัณฑิต ภาควิชาวัดผลและวิจัยการ
ศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

ภาษาต่างประเทศ

- Calfee, R., & Perfumo, P. (1996). **Writing portfolios in the classroom**. New Jersey: Lawrence Erlbaum Associates, Inc.
- Cole, J.D., Ryan, W.C., Kick, F., & Mathies, K.B. (2000). **Portfolio across the curriculum and beyond**. California: Corwin Press, Inc.
- Jardine, A.S. (1996). Key points of the "Authentic Assessment portfolio". **Journal of Intervention in School and Clinic** 31: 252-253
- Judith, A. (2000). Rubrics, scoring guide, and performance criteria: classroom tool for assessment and improving student learning. **ERIC Document Reproduction Service No. ED 446100**. Paper Present at the Annual Meeting of the American Educational Research Association.
- Goodrich, H. (1997). **Understanding rubrics**. [Online]. Available: <http://cdnet2.car.chula.ac.th/pdfhtml> [2000, July. 18]
- Kevin, C., & Rhoton, H. (1998). A comparison of two scoring strategies for performance assessment. **ERIC Document Reproduction Service No. ED 422397**. Paper Present at the Annual Meeting of the National Council on Measurement in Educational.
- Kristie, W., Andrea, K., & Gina, K. (1998). Alternative approaches to scoring: the effects of using different scoring methods on the validity of scores from a performance assessment. **ERIC Document Reproduction Service No. ED 422397**.
- Kubiszyn, T., & Borich, G. (2000). **Educational testing and measurement: classroom application and practice**. New York: John Wiley & Sons. Inc.
- Lane, S and others. (1996). Generalizability and validity of a mathematics performance assessment. **Journal of Educational Measurement**. 33: 71-91
- Moore, A. D., & Young, Z. (1997). Clarifying the blurred image: estimating the inter-rater reliability of performance assessment. **ERIC Document Reproduction Service No. ED 414319**.
- Moskal, M. (2000). Scoring rubrics part I: what and when. **ERIC Document Reproduction Service No. ED 446110**.
- Moskal, M. (2000). Scoring rubrics part II: what and when. **ERIC Document Reproduction Service No. ED 446111**.

- Nitko, A.J. (1996). **Educational assessment of students**. Englewood Cliffs, New Jersey: Second edition. Prentice-Hall.
- Ronis, D. (2000). **Brain-compatible assessments**. New York: Skylight Training and Publishing Inc.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ก
เครื่องมือที่ใช้ในการวิจัย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

เกณฑ์การให้คะแนนแบบภาพรวม
(การเขียนเรื่องประกอบภาพ)

| ระดับ คะแนน | ลักษณะของผลงาน |
|--------------------|---|
| 4 (ดีมาก) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - มีความรู้ในเรื่องที่เขียนดีมาก เขียนได้ใจความน่าสนใจ มีความคิดสร้างสรรค์ดีมาก - เนื้อหามีความสัมพันธ์กับภาพประกอบและชื่อเรื่อง - เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความ กะทัดรัด และมีความต่อเนื่องดีมาก - ใช้ถ้อยคำ สำนวนภาษาได้อย่างถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อความหมายได้ชัดเจนดีมาก - ไม่มีข้อผิดพลาดเลยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ |
| 3 (ดี) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - มีความรู้ในเรื่องที่เขียนดี เขียนได้ใจความน่าสนใจ มีความคิดสร้างสรรค์ดี เนื้อหาส่วนใหญ่มีความสัมพันธ์กับภาพประกอบและชื่อเรื่อง - เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความ กะทัดรัด และมีความต่อเนื่องดี - มีข้อผิดพลาดเล็กน้อยในการใช้ถ้อยคำ สำนวนภาษา สามารถสื่อความหมายได้ไม่ทำให้ความหมายเปลี่ยน - มีข้อผิดพลาดเล็กน้อยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 10 % แต่ไม่เกิน 50 % |
| 2 (พอใช้) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - มีความรู้ในเรื่องที่เขียนพอสมควร เขียนได้ใจความพอสมควร ขาดความคิดสร้างสรรค์ เนื้อหาบางส่วนไม่สัมพันธ์กับภาพประกอบและชื่อเรื่อง - เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความ กะทัดรัด และมีความต่อเนื่องค่อนข้างน้อย - ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้อง และไม่เหมาะสม บางครั้งไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดค่อนข้างมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - มีความรู้ในเรื่องที่เขียนน้อย เขียนเนื้อหาไม่ตรงกับชื่อเรื่อง ไม่น่าสนใจ ไม่ได้ใจความ ขาดความคิดสร้างสรรค์ สำนวนภาพประกอบและชื่อเรื่องไม่สัมพันธ์กับเนื้อหา - ไม่สามารถเรียบเรียงความคิดอย่างสมเหตุสมผล และไม่มีมีความต่อเนื่อง - ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้องและไม่เหมาะสม ไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 90 % |

เกณฑ์การให้คะแนนแบบภาพรวม
(การบันทึกความรู้เรื่องี่อ่าน)

| ระดับ คะแนน | ลักษณะของผลงาน |
|--------------------|---|
| 4 (ดีมาก) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - ระบุประเด็นสำคัญของเรื่องี่อ่านได้ดี 80-100% ตามรูปแบบที่กำหนด - การลำดับเนื้อหาหลักและรายละเอียดย่อยสัมพันธ์กันตลอดทั้งเรื่องี่อ่าน - ใช้ถ้อยคำ สำนวนภาษาได้อย่างถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อความหมายได้ชัดเจนดีมาก - ไม่มีข้อผิดพลาดเลยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ |
| 3 (ดี) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - ระบุประเด็นสำคัญของเรื่องี่อ่านได้ 71-79% ตามรูปแบบที่กำหนด - การลำดับเนื้อหาหลักสัมพันธ์กันดี แต่มีข้อผิดพลาดเล็กน้อยในการลำดับรายละเอียดย่อย - มีข้อผิดพลาดเล็กน้อยในการใช้ถ้อยคำ สำนวนภาษา สามารถสื่อความหมายได้ไม่ทำให้ความหมายเปลี่ยน - มีข้อผิดพลาดเล็กน้อยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 10 % แต่ไม่เกิน 50 % |
| 2 (พอใช้) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - ระบุประเด็นสำคัญของเรื่องี่อ่านได้ 51-70 % ตามรูปแบบที่กำหนด - มีข้อผิดพลาดค่อนข้างมากในการลำดับเนื้อหาหลักและรายละเอียดย่อยให้สัมพันธ์กันตลอดเรื่องี่อ่าน - ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้อง และไม่เหมาะสม บางครั้งไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดค่อนข้างมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - ไม่สามารถระบุประเด็นสำคัญของเรื่องี่อ่านได้ - ไม่มีการลำดับเนื้อหาหลักและรายละเอียดย่อยได้สัมพันธ์กัน - ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้องและไม่เหมาะสม ไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 90 % |

เกณฑ์การให้คะแนนแบบภาพรวม
(การเขียนจดหมาย)

| ระดับ คะแนน | ลักษณะของผลงาน |
|--------------------|---|
| 4 (ดีมาก) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - แบบฟอร์มของการเขียนจดหมาย และการจำหน่ายของจดหมาย ถูกต้อง - เขียนเนื้อหาชัดเจน กระชับรัด ได้ใจความ ตรงประเด็น 80-100 % โดยลำดับความคิด เหตุการณ์อย่างต่อเนื่อง - ใช้ถ้อยคำ สำนวนภาษาได้อย่างถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อความหมายได้ชัดเจนดีมาก - ไม่มีข้อผิดพลาดเลยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ |
| 3 (ดี) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - มีข้อผิดพลาดเล็กน้อย ในเรื่องแบบฟอร์มของการเขียนจดหมาย และการจำหน่ายของจดหมาย - เขียนเนื้อหาชัดเจน ตรงประเด็น 75-80 % แต่มีข้อบกพร่องเล็กน้อยในการลำดับความคิด เหตุการณ์ของเรื่อง - มีข้อผิดพลาดเล็กน้อยในการใช้ถ้อยคำ สำนวนภาษา สามารถสื่อความหมายได้ไม่ทำให้ความหมายเปลี่ยน - มีข้อผิดพลาดเล็กน้อยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 10 % แต่ไม่เกิน 50 % |
| 2 (พอใช้) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - มีข้อผิดพลาดค่อนข้างมาก ในเรื่องแบบฟอร์มของการเขียนจดหมาย และการจำหน่ายของจดหมาย - เขียนเนื้อหาชัดเจน ตรงประเด็น 50-74% มีข้อบกพร่องค่อนข้างมาก ในการลำดับความคิด เหตุการณ์ของเรื่อง - ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้อง และไม่เหมาะสม บางครั้งไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดค่อนข้างมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | <p>ผลงานมีลักษณะตามเกณฑ์ดังนี้</p> <ul style="list-style-type: none"> - แบบฟอร์มของการเขียนจดหมาย และการจำหน่ายของจดหมายไม่ถูกต้อง - เขียนเนื้อหาไม่ชัดเจน ไม่ตรงประเด็น การลำดับความคิด เหตุการณ์ของเรื่องไม่ต่อเนื่อง - ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้องและไม่เหมาะสม ไม่สามารถสื่อความหมายให้เข้าใจได้ - มีข้อผิดพลาดมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 90 % |

เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ
การเขียนเรื่องประกอบภาพ

| ระดับ คะแนน | เนื้อหา | การลำดับเรื่อง | การใช้ภาษา | กลไกการเขียน |
|--------------------|--|--|--|---|
| 4 (ดีมาก) | มีความรู้ในเรื่องที่เขียนดีมาก เขียนได้ใจความ น่าสนใจ มีความคิดสร้างสรรค์ดีมาก เนื้อหามีความสัมพันธ์กับภาพประกอบและชื่อเรื่อง | เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความกะทัดรัด และมีความต่อเนื่องดีมาก | ใช้ถ้อยคำ สำนวนภาษาได้อย่าง ถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อความหมายได้ชัดเจนดีมาก | ไม่มีข้อผิดพลาดเลยในการสะกดคำและ การใช้เครื่องหมายวรรคตอนต่างๆ |
| 3 (ดี) | มีความรู้ในเรื่องที่เขียนดี เขียนได้ใจความ น่าสนใจ มีความคิดสร้างสรรค์ดี เนื้อหาส่วนใหญ่มีความสัมพันธ์กับภาพประกอบและชื่อเรื่อง | เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความกะทัดรัด และมีความต่อเนื่องดี | มีข้อผิดพลาดเล็กน้อยในการใช้ถ้อยคำ สำนวนภาษา สามารถสื่อความหมายได้ ไม่ทำให้ความหมายเปลี่ยน | มีข้อผิดพลาดเล็กน้อยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 10 % แต่ไม่เกิน 50 % |
| 2 (พอใช้) | มีความรู้ในเรื่องที่เขียนพอสมควร เขียนได้ใจความพอสมควร ขาดความคิดสร้างสรรค์ เนื้อหามีบางส่วนสัมพันธ์กับภาพประกอบและชื่อเรื่อง | เรียบเรียงความคิดอย่างสมเหตุสมผล เขียนได้ความกะทัดรัด และมีความต่อเนื่องค่อนข้างน้อย | ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้องและไม่เหมาะสม บางครั้งไม่สามารถสื่อความหมายให้เข้าใจได้ | มีข้อผิดพลาดค่อนข้างมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | มีความรู้ในเรื่องที่เขียนน้อย เขียนเนื้อหาไม่ตรงกับชื่อเรื่อง ไม่น่าสนใจ ไม่ได้ใจความ ขาดความคิดสร้างสรรค์ ส่วนภาพประกอบและชื่อเรื่องไม่สัมพันธ์กับเนื้อหา | ไม่สามารถเรียบเรียงความคิดอย่างสมเหตุสมผลและ ไม่มีความต่อเนื่อง | ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้องและไม่เหมาะสม ไม่สามารถสื่อความหมายให้เข้าใจได้ | มีข้อผิดพลาดมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 90 % |

เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ
(การบันทึกความรู้เรื่องที่อ่าน)

| ระดับคะแนน | การระบุประเด็นสำคัญของเรื่อง ที่อ่านได้ | การลำดับเรื่อง | การใช้ภาษา | กลไกการเขียน |
|--------------------|--|---|---|--|
| 4 (ดีมาก) | ระบุประเด็นสำคัญของเรื่อง ที่อ่านได้ 80-100% ตามรูปแบบที่กำหนด | การลำดับเนื้อหาหลักและรายละเอียดย่อยสัมพันธ์กันดีตลอดเรื่อง ที่อ่าน | ใช้ถ้อยคำ สำนวนภาษาได้อย่าง ถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อความหมายได้ชัดเจนดีมาก | ไม่มีข้อผิดพลาดเลยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ |
| 3 (ดี) | ระบุประเด็นสำคัญของเรื่อง ที่อ่านได้ 71-79% ตามรูปแบบที่กำหนด | การลำดับเนื้อหาหลักสัมพันธ์กันดี แต่มีข้อผิดพลาดเล็กน้อยในการลำดับ รายละเอียดย่อย | มีข้อผิดพลาดเล็กน้อยในการใช้ถ้อยคำ สำนวนภาษา สามารถสื่อความหมายได้ ไม่ทำให้ความหมายเปลี่ยน | มีข้อผิดพลาดเล็กน้อยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆเกิน 10 % แต่ไม่เกิน 50 % |
| 2 (พอใช้) | ระบุประเด็นสำคัญของเรื่อง ที่อ่าน ได้ 51-70% ตามรูปแบบที่กำหนด | มีข้อผิดพลาดค่อนข้างมากในการลำดับเนื้อหาหลักและรายละเอียดย่อยให้สัมพันธ์กันตลอดเรื่อง ที่อ่าน | ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้องและม่เหมาะสม บางครั้งไม่สามารถสื่อความหมายให้เข้าใจได้ | มีข้อผิดพลาดค่อนข้างมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | ไม่สามารถระบุประเด็นสำคัญของเรื่อง ที่อ่านได้ | ไม่มีการลำดับเนื้อหาหลักและรายละเอียดย่อยให้สัมพันธ์กัน | ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้อง และไม่เหมาะสม ไม่สามารถสื่อความหมายให้เข้าใจได้ | มีข้อผิดพลาดมากใน การสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 90 % |

เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ
(การเขียนจดหมาย)

| ระดับ คะแนน | แบบฟอร์มของจดหมาย | เนื้อหา | การใช้ภาษา | กลไกการเขียน |
|--------------------|--|---|--|---|
| 4 (ดีมาก) | แบบฟอร์มของการเขียนจดหมายและการจำหน่ายของจดหมายถูกต้อง | เขียนเนื้อหาชัดเจน กะทัดรัด ได้ใจความ รงประเด็น 80-100% โดยลำดับความคิด เหตุการณ์อย่างต่อเนื่อง | ใช้ถ้อยคำ สำนวนภาษาได้อย่างถูกต้อง และเหมาะสม เข้าใจง่าย สามารถสื่อความหมายได้ชัดเจนดีมาก | ไม่มีข้อผิดพลาดเลยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ |
| 3 (ดี) | มีข้อผิดพลาดเล็กน้อยในเรื่องแบบฟอร์มของการเขียนจดหมายและการจำหน่ายของจดหมาย | เขียนเนื้อหาชัดเจน ตรงประเด็น 75-80 % แต่มีข้อบกพร่องเล็กน้อยในการลำดับความคิด เหตุการณ์ของเรื่อง | มีข้อผิดพลาดเล็กน้อยในการใช้ถ้อยคำ สำนวนภาษา สามารถสื่อความหมายได้ ไม่ทำให้ความหมายเปลี่ยน | มีข้อผิดพลาดเล็กน้อยในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 10 % แต่ไม่เกิน 50 % |
| 2 (พอใช้) | มีข้อผิดพลาดค่อนข้างมากในเรื่องแบบฟอร์มของการเขียนจดหมายและการจำหน่ายของจดหมาย | เขียนเนื้อหาได้ตรง ประเด็น 50-74% มีข้อบกพร่องค่อนข้างมาก ในการลำดับความคิด เหตุการณ์ของเรื่อง | ใช้ถ้อยคำ สำนวนภาษา บางส่วนไม่ถูกต้องและไม่เหมาะสม บางครั้งไม่สามารถสื่อความหมายให้เข้าใจได้ | มีข้อผิดพลาดค่อนข้างมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 50 % |
| 1 (ควรปรับปรุง) | แบบฟอร์มของการเขียนจดหมายและการจำหน่ายของจดหมายไม่ถูกต้อง | เขียนเนื้อหาไม่ชัดเจน ไม่ตรงประเด็น การลำดับความคิด เหตุการณ์ของเรื่องไม่ต่อเนื่อง | ใช้ถ้อยคำ สำนวนภาษา ไม่ถูกต้อง และไม่เหมาะสม ไม่สามารถสื่อความหมายให้เข้าใจได้ | มีข้อผิดพลาดมากในการสะกดคำและการใช้เครื่องหมายวรรคตอนต่างๆ เกิน 90 % |



ภาคผนวก ข
ตัวอย่างผลงานนักเรียน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ชิ้นงานที่เลือกมาตีพิมพ์ในหนังสือ

ในครั้งที่ ๓

วิชา ภาษาไทย

ชื่อ การเขียนเรื่องประกอบฉาก

กิจกรรม ครูช่วยนักเรียน

<1๘.๓๔

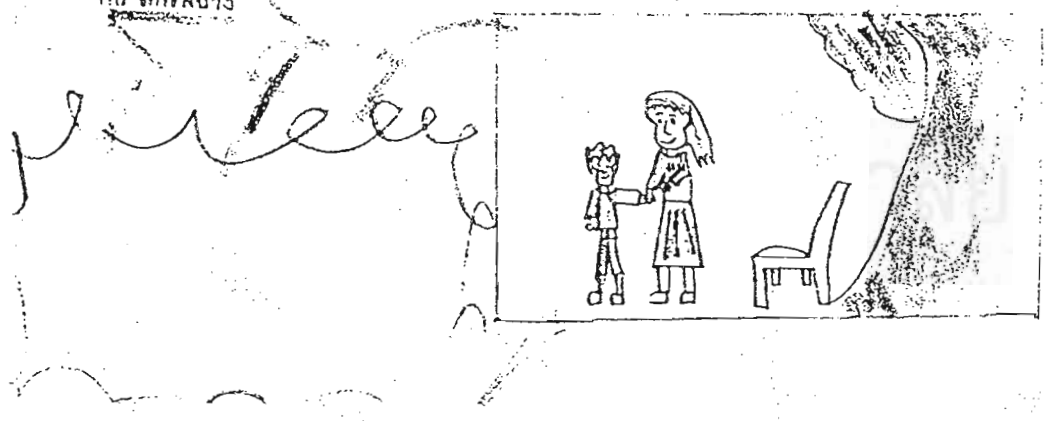
ชื่อ อ.อ.ศิริรัตน์

อุดมการณ์ ชั้น ๕/๕

ครั้งที่ ๑๖

เรื่อง พระอุกขตงแม่

แม่มีสี่พระอุกขตงมากมาย แม่ให้กำหนัดรดที่ขงอุรวและแม่ให้
 ก็ใช้เวลามุ่งรกรวดหา หาตงเป็นลูกที่ติดแม่ โดยการให้ฟิ่งแม่ ยี่สิบ
 ยี่สิบหกลูก ไม่เอาแม่ให้แม่ คนบางตัวแม่ก็ไปทิวี่สิบดูของ
 ใดเข้าเป็นพี่ใหญ่ และก็มีแม่พิมพ์แม่ที่ราคา เขาก็ตั้งลดจนดูแม่ หักแม่ไปมาก
 ที่คิดในการทดแทนแม่หรือลูก ก็แม่ที่ขงหาตั้งแม่อีกสิบโต และหาตงอุกขตงแม่ตัวอื่น
 ไปดกก็ให้แม่อยู่ใกล้แม่ที่ขงหรือ เขาตงกตั้งแม่แม่ ถ้าใจนั้นไม่ดแม่แม่จะพหแม่
 คนและนี้หตงแม่แม่ หตงนี้เขาก็ตงและได้ความดีกับแม่ที่ดตงแม่คนและตงแม่แม่
 คำแม่นี้ไม่เท่าไรมาก เขาก็จะใช้ตงแม่และตงแม่ที่ทำแม่ ถ้าใครดแม่แม่เป็นตงแม่ก็
 แม่ คนคนนั้นตงแม่ที่ตงหรือ และตงแม่ได้แม่ แม่เป็นแม่ให้ ดั่งนี้เขาก็คงตงแม่แม่หรือคนใจ
 กษให้แม่ขว



โรงเรียนเทศบาลเมือง
โบราณคดี

วิชา ภาษาไทย เรื่อง การเขียนบันทึก
กิจกรรม บันทึกประจำวัน เรื่อง
ชื่อ ด.ช. ณัฏฐพงษ์ เปี้ยทอง ชั้น ป.๕/๕ เลขที่ ๑๑

วันที่เขียน

บันทึกการอ่าน

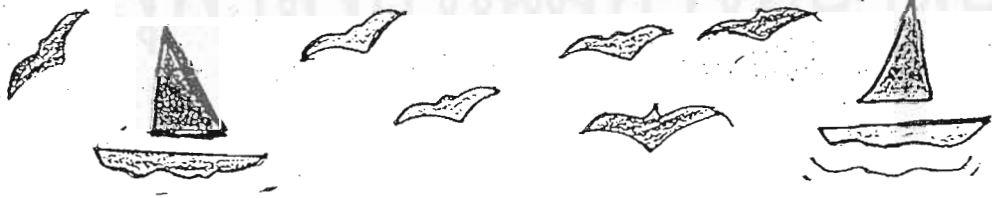
ชื่อเรื่อง นกกับนกกรงศา ผู้แต่ง นายฉลอง ยิ้มโสมบรรณ

ณ บึงใหญ่ แห่งหนึ่ง ใจกลางป่า มีน้ำ และนกกรงศา คู่หนึ่ง
กำลังเล่นน้ำอย่างสนุกสนาน ปลาเล็กปลาน้อย กินเป็น อาหาร โดยไม่รู้ตัว เลย
ว่า ภายในบึงใหญ่มีปลา และ นกกรงศา ทั้งสอง ฝ่าย พยายาม ฆ่ากันเอง ยิ่งได้ปลา
มากเท่าละก็ พยายาม อยู่ ว่านกกรงศา กับ ปลา ก็อย่างไรมิใช่ ในบึง
ซึ่งนกกรงศา ค่อยๆ กิน อย่าง ไม่รีบร้อนนัก เมื่อวันหนึ่ง นกกรงศา เห็น
ปลาตัวหนึ่ง ว่า "นี่น่า กินยากขนาดนี้ รวดเร็วไม่ได้อะไร" แต่อย่าง
ไม่ว่า นกกรงศา ก็ยังคิดกิน ปลาอย่าง ไม่คิดถึงใครทั้งนั้น ส่วนนายพรวน
พบเห็นนกกรงศา และ ปลา เมล็ด เขาก็ ยกมือ ขึ้นมา ทำ จิตใจ ก็
นกกรงศา เห็นไป เห็นเข้า พอดี เลย ร้อง โอดครวญ ไปได้ ในบึงที่
นั่น นกกรงศา เห็น นายพรวน เห็นกัน แต่อย่าง ก็ไม่รีบร้อน สบายใจ
คืนหนึ่ง ของมันในบึง ทำให้มันไม่กล้า มากรวดเร็ว มันได้ มันเจ็ด ตก เป็น
อาหารของนายพรวน ในที่สุด

ตัวละครที่ชอบ นกกรงศา
สาเหตุที่ชอบ สบายใจ สบายใจ

ข้อคิดที่ได้จากเรื่อง นายพรวน
สาเหตุที่ไม่ชอบ เพราะได้จับมาเป็นอาหาร

ข้อคิดที่ได้จากเรื่อง ไม่ชอบ ฝรั่งค้าขายที่หัวรั้วตี ซื่อสัตย์ อยู่ในศีลธรรม



พินิจที่เลื่อมเจ้า
โรงเรียนไพทอตนศึกษา
ใบงานที่ ๒

วิชา ภาษาไทย เรื่อง กาวเขียนจดหมายส่วนตัว
กิจกรรม ไปรษณีย์สี่สาว ๑๓ ก.ค. ๕๕

ชื่อ ด.ญ. กรองกาญจน์ ชัยปิตินานนท์ ชั้น ๒/๕ เลขที่ ๒๖



๑๒๒/๖๖๓ ถนนวิภาวดีรังสิต
แขวง กิ่งสองห้อง เขตหลักสี่
กรุงเทพฯ ๑๐๒๑๐

๑๓ กรกฎาคม ๒๕๕๕

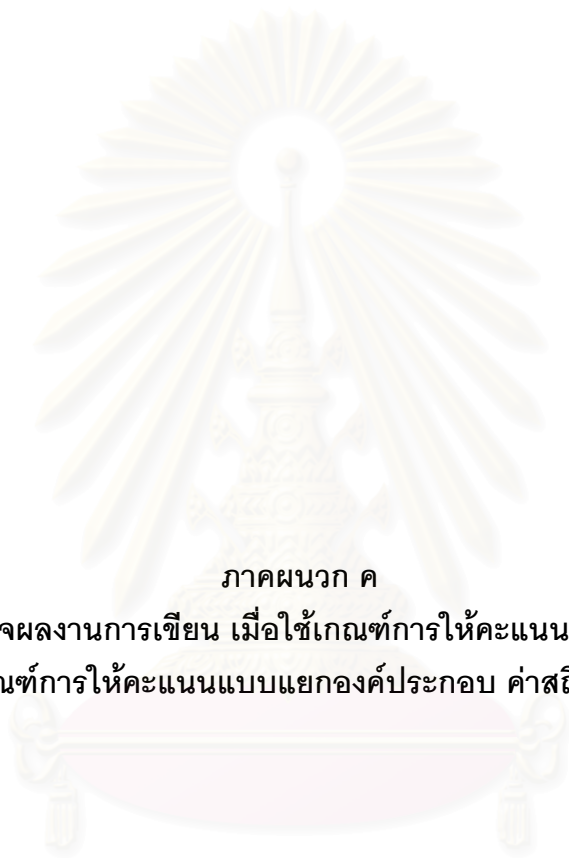
เกตุเพื่อนรัก

จดหมายฉบับนี้เขียนขึ้นเวลาที่ฉันเขียนกิจกรรม แอ่ยบาย
น้องแก้ว เธอมีเพื่อนใหม่หรือยัง ส่วนฉันมีเพื่อนใหม่มากมาย
ฉันสบายดีส่วนเธอสบายดีหรือเปล่า โรงเรียนของฉันมีกิจกรรม
มากมาย งานที่เรียนอยู่ก็มาก เช่น วันสุนทราภรณ์ ละครกิจกรรม
ในหน้าเรียน ฉันก็ได้ไปร่วมกิจกรรมได้วันความร่วมมากมาย
เพราะที่นี่จัดไม่เหมือนที่ก่อนๆ ที่โรงเรียนของเธอจะมีกิจกรรม
อะไรบ้าง เขียนจดหมายเล่าให้ฉันฟังบ้างนะ

สุดท้ายนี้ ฉันขอให้คุณมีความสุขเป็นที่ยาวนาน
คุณแม่ตลอดไป

รักและคิดถึง

กรองกาญจน์ ชัยปิตินานนท์



ภาคผนวก ค

ผลการตรวจผลงานการเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม
และเกณฑ์การให้คะแนนแบบแยกองค์ประกอบ ค่าสถิติพื้นฐาน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 12 ผลการตรวจผลงานการเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม
และ ค่าสถิติพื้นฐาน

| คนที่ | ผู้ประเมินคนที่ 1 | | | | ผู้ประเมินคนที่ 2 | | | | ผู้ประเมินคนที่ 3 | | | | ผู้ประเมินคนที่ 4 | | | |
|-------|-------------------|-----------|-----------|-----|-------------------|-----------|-----------|-----|-------------------|-----------|-----------|-----|-------------------|-----------|-----------|-----|
| | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม |
| 1 | 4 | 4 | 3 | 11 | 3 | 2 | 2 | 7 | 3 | 3 | 4 | 10 | 3 | 3 | 3 | 9 |
| 2 | 4 | 3 | 4 | 11 | 2 | 3 | 3 | 8 | 3 | 3 | 3 | 9 | 3 | 2 | 3 | 8 |
| 3 | 4 | 4 | 3 | 11 | 3 | 2 | 2 | 7 | 4 | 3 | 3 | 10 | 3 | 3 | 2 | 8 |
| 4 | 4 | 4 | 4 | 12 | 3 | 2 | 4 | 9 | 4 | 4 | 4 | 12 | 4 | 3 | 4 | 11 |
| 5 | 3 | 3 | 4 | 10 | 2 | 2 | 2 | 6 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| 6 | 4 | 4 | 3 | 11 | 2 | 2 | 3 | 7 | 3 | 4 | 3 | 10 | 4 | 4 | 3 | 11 |
| 7 | 4 | 4 | 3 | 11 | 2 | 3 | 2 | 7 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| 8 | 4 | 3 | 4 | 11 | 2 | 3 | 2 | 7 | 3 | 3 | 3 | 9 | 4 | 3 | 3 | 10 |
| 9 | 4 | 4 | 4 | 12 | 3 | 2 | 2 | 7 | 3 | 3 | 3 | 9 | 3 | 3 | 4 | 10 |
| 10 | 4 | 4 | 4 | 12 | 2 | 2 | 3 | 7 | 3 | 3 | 4 | 10 | 4 | 4 | 4 | 12 |
| 11 | 3 | 4 | 3 | 10 | 2 | 3 | 4 | 9 | 3 | 3 | 3 | 9 | 4 | 4 | 3 | 11 |
| 12 | 4 | 4 | 3 | 11 | 2 | 2 | 2 | 6 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| 13 | 3 | 4 | 3 | 10 | 2 | 3 | 2 | 7 | 3 | 3 | 4 | 10 | 4 | 3 | 3 | 10 |
| 14 | 3 | 4 | 3 | 10 | 2 | 3 | 2 | 7 | 3 | 3 | 4 | 10 | 3 | 3 | 3 | 9 |
| 15 | 4 | 4 | 4 | 12 | 3 | 3 | 2 | 8 | 4 | 3 | 3 | 10 | 4 | 4 | 3 | 11 |
| 16 | 3 | 4 | 4 | 11 | 3 | 3 | 3 | 9 | 4 | 3 | 4 | 11 | 3 | 3 | 4 | 10 |
| 17 | 4 | 3 | 4 | 11 | 2 | 2 | 1 | 5 | 4 | 3 | 3 | 10 | 3 | 3 | 3 | 9 |
| 18 | 4 | 4 | 4 | 12 | 2 | 2 | 3 | 7 | 3 | 3 | 4 | 10 | 4 | 4 | 3 | 11 |
| 19 | 4 | 4 | 4 | 12 | 2 | 2 | 3 | 7 | 3 | 3 | 3 | 9 | 2 | 4 | 4 | 10 |
| 20 | 3 | 3 | 3 | 9 | 1 | 2 | 2 | 5 | 3 | 4 | 4 | 11 | 3 | 4 | 2 | 9 |
| 21 | 4 | 4 | 3 | 11 | 3 | 2 | 1 | 6 | 3 | 4 | 3 | 10 | 3 | 4 | 4 | 11 |
| 22 | 4 | 4 | 4 | 12 | 2 | 3 | 4 | 9 | 4 | 4 | 4 | 12 | 4 | 4 | 4 | 12 |
| 23 | 3 | 4 | 3 | 10 | 2 | 3 | 2 | 7 | 3 | 3 | 3 | 9 | 3 | 4 | 4 | 11 |
| 24 | 3 | 3 | 3 | 9 | 1 | 2 | 3 | 6 | 3 | 3 | 4 | 10 | 3 | 3 | 3 | 9 |
| 25 | 4 | 4 | 4 | 12 | 3 | 3 | 4 | 10 | 4 | 3 | 4 | 11 | 4 | 3 | 4 | 11 |
| 26 | 4 | 4 | 4 | 12 | 4 | 3 | 3 | 10 | 3 | 4 | 4 | 11 | 4 | 4 | 4 | 12 |
| 27 | 4 | 4 | 3 | 11 | 2 | 4 | 3 | 9 | 4 | 4 | 3 | 11 | 4 | 4 | 3 | 11 |
| 28 | 4 | 3 | 4 | 11 | 2 | 2 | 2 | 6 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |

ตาราง 12 (ต่อ) ผลการตรวจผลงานการเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบภาพรวม
และ ค่าสถิติพื้นฐาน

| คนที่ | ผู้ประเมินคนที่ 1 | | | | ผู้ประเมินคนที่ 2 | | | | ผู้ประเมินคนที่ 3 | | | | ผู้ประเมินคนที่ 4 | | | |
|-----------|-------------------|-----------|-----------|------|-------------------|-----------|-----------|-------|-------------------|-----------|-----------|------|-------------------|-----------|-----------|-------|
| | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม |
| 29 | 4 | 4 | 4 | 12 | 2 | 2 | 3 | 7 | 3 | 4 | 3 | 10 | 3 | 4 | 4 | 11 |
| 30 | 4 | 4 | 4 | 12 | 3 | 2 | 3 | 8 | 3 | 3 | 4 | 10 | 4 | 3 | 4 | 11 |
| 31 | 4 | 3 | 4 | 11 | 2 | 2 | 3 | 7 | 3 | 4 | 3 | 10 | 3 | 3 | 4 | 10 |
| 32 | 4 | 4 | 3 | 11 | 3 | 3 | 3 | 9 | 3 | 3 | 4 | 10 | 3 | 4 | 3 | 10 |
| 33 | 4 | 4 | 4 | 12 | 2 | 3 | 3 | 8 | 3 | 4 | 4 | 11 | 4 | 4 | 4 | 12 |
| 34 | 3 | 4 | 3 | 10 | 2 | 3 | 3 | 8 | 3 | 3 | 3 | 9 | 4 | 3 | 3 | 10 |
| 35 | 4 | 4 | 4 | 12 | 3 | 2 | 2 | 7 | 4 | 3 | 4 | 11 | 4 | 3 | 3 | 10 |
| 36 | 4 | 4 | 4 | 12 | 2 | 3 | 3 | 8 | 3 | 4 | 3 | 10 | 4 | 3 | 4 | 11 |
| 37 | 3 | 4 | 4 | 11 | 2 | 2 | 2 | 6 | 4 | 3 | 3 | 10 | 3 | 4 | 3 | 10 |
| 38 | 4 | 4 | 3 | 11 | 2 | 2 | 2 | 6 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| 39 | 4 | 4 | 4 | 12 | 2 | 2 | 2 | 6 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| \bar{X} | 3.74 | 3.79 | 3.59 | 11.1 | 2.28 | 2.46 | 2.56 | 7.308 | 3.26 | 3.28 | 3.41 | 9.95 | 3.41 | 3.385 | 3.33 | 10.13 |
| S.D. | 0.44 | 0.41 | 0.498 | 0.86 | 0.6 | 0.55 | 0.75 | 1.26 | 0.44 | 0.46 | 0.5 | 0.86 | 0.55 | 0.544 | 0.58 | 1.105 |

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย


ตาราง 13 ผลการตรวจผลงานการเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ
และ ค่าสถิติพื้นฐาน

| คนที่ | ผู้ประเมินคนที่ 1 | | | | ผู้ประเมินคนที่ 2 | | | | ผู้ประเมินคนที่ 3 | | | | ผู้ประเมินคนที่ 4 | | | |
|-------|-------------------|-----------|-----------|-----|-------------------|-----------|-----------|-----|-------------------|-----------|-----------|-----|-------------------|-----------|-----------|-----|
| | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม |
| 1 | 3 | 3 | 2 | 8 | 3 | 2 | 2 | 7 | 3 | 3 | 2 | 8 | 3 | 2 | 2 | 7 |
| 2 | 3 | 2 | 2 | 7 | 3 | 2 | 2 | 7 | 2 | 3 | 3 | 8 | 2 | 2 | 2 | 6 |
| 3 | 2 | 3 | 2 | 7 | 2 | 2 | 2 | 6 | 3 | 3 | 3 | 9 | 3 | 2 | 2 | 7 |
| 4 | 2 | 3 | 3 | 8 | 3 | 3 | 2 | 8 | 3 | 4 | 4 | 11 | 4 | 3 | 4 | 11 |
| 5 | 2 | 3 | 3 | 8 | 2 | 3 | 2 | 7 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| 6 | 3 | 4 | 3 | 10 | 3 | 3 | 2 | 8 | 3 | 2 | 3 | 8 | 3 | 3 | 3 | 9 |
| 7 | 2 | 4 | 3 | 9 | 2 | 3 | 3 | 8 | 2 | 3 | 2 | 7 | 3 | 4 | 4 | 11 |
| 8 | 2 | 3 | 3 | 8 | 2 | 2 | 2 | 6 | 3 | 3 | 3 | 9 | 3 | 4 | 3 | 10 |
| 9 | 4 | 2 | 2 | 8 | 3 | 2 | 2 | 7 | 3 | 3 | 3 | 9 | 2 | 3 | 4 | 9 |
| 10 | 4 | 3 | 4 | 11 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 | 4 | 4 | 3 | 11 |
| 11 | 1 | 4 | 4 | 9 | 2 | 4 | 3 | 9 | 3 | 3 | 3 | 9 | 2 | 4 | 3 | 9 |
| 12 | 2 | 3 | 2 | 7 | 2 | 3 | 2 | 7 | 3 | 3 | 3 | 9 | 2 | 2 | 4 | 8 |
| 13 | 3 | 2 | 2 | 7 | 3 | 3 | 2 | 8 | 3 | 3 | 3 | 9 | 4 | 2 | 2 | 8 |
| 14 | 2 | 2 | 2 | 6 | 2 | 2 | 2 | 6 | 3 | 2 | 2 | 7 | 2 | 2 | 3 | 7 |
| 15 | 2 | 2 | 3 | 7 | 2 | 3 | 3 | 8 | 3 | 3 | 3 | 9 | 3 | 4 | 3 | 10 |
| 16 | 4 | 4 | 3 | 11 | 2 | 2 | 3 | 7 | 3 | 4 | 3 | 10 | 4 | 3 | 4 | 11 |
| 17 | 2 | 2 | 2 | 6 | 2 | 2 | 2 | 6 | 3 | 2 | 2 | 7 | 3 | 2 | 2 | 7 |
| 18 | 4 | 4 | 2 | 10 | 4 | 4 | 4 | 12 | 3 | 3 | 3 | 9 | 4 | 4 | 3 | 11 |
| 19 | 4 | 4 | 4 | 12 | 3 | 3 | 4 | 10 | 3 | 2 | 3 | 8 | 3 | 3 | 4 | 10 |
| 20 | 4 | 2 | 3 | 9 | 3 | 3 | 3 | 9 | 2 | 3 | 3 | 8 | 3 | 3 | 4 | 10 |
| 21 | 3 | 4 | 4 | 11 | 3 | 3 | 4 | 10 | 3 | 4 | 3 | 10 | 4 | 4 | 3 | 11 |
| 22 | 4 | 4 | 4 | 12 | 3 | 3 | 4 | 10 | 3 | 3 | 2 | 8 | 3 | 4 | 4 | 11 |
| 23 | 3 | 3 | 3 | 9 | 3 | 2 | 3 | 8 | 3 | 3 | 3 | 9 | 4 | 3 | 2 | 9 |
| 24 | 3 | 3 | 3 | 9 | 3 | 2 | 2 | 7 | 3 | 3 | 2 | 8 | 2 | 3 | 3 | 8 |
| 25 | 3 | 2 | 2 | 7 | 2 | 2 | 2 | 6 | 3 | 3 | 2 | 8 | 4 | 3 | 4 | 11 |
| 26 | 3 | 4 | 4 | 11 | 4 | 4 | 4 | 12 | 2 | 2 | 3 | 7 | 4 | 4 | 4 | 12 |
| 27 | 4 | 4 | 3 | 11 | 4 | 4 | 4 | 12 | 3 | 3 | 3 | 9 | 3 | 4 | 2 | 9 |
| 28 | 3 | 3 | 4 | 10 | 3 | 3 | 3 | 9 | 3 | 4 | 3 | 10 | 3 | 3 | 4 | 10 |

ตาราง 13 (ต่อ) ผลการตรวจผลงานการเขียน เมื่อใช้เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ
และ ค่าสถิติพื้นฐาน

| คนที่ | ผู้ประเมินคนที่ 1 | | | | ผู้ประเมินคนที่ 2 | | | | ผู้ประเมินคนที่ 3 | | | | ผู้ประเมินคนที่ 4 | | | |
|-----------|-------------------|-----------|-----------|------|-------------------|-----------|-----------|-------|-------------------|-----------|-----------|------|-------------------|-----------|-----------|-------|
| | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม | ชั้นที่ 1 | ชั้นที่ 2 | ชั้นที่ 3 | รวม |
| 29 | 3 | 4 | 4 | 11 | 3 | 4 | 4 | 11 | 3 | 3 | 3 | 9 | 3 | 4 | 3 | 10 |
| 30 | 3 | 3 | 4 | 10 | 3 | 3 | 4 | 10 | 3 | 3 | 3 | 9 | 4 | 3 | 4 | 11 |
| 31 | 4 | 3 | 4 | 11 | 3 | 3 | 3 | 9 | 4 | 4 | 3 | 11 | 4 | 3 | 4 | 11 |
| 32 | 3 | 2 | 2 | 7 | 3 | 3 | 2 | 8 | 3 | 4 | 4 | 11 | 3 | 4 | 4 | 11 |
| 33 | 4 | 4 | 4 | 12 | 3 | 4 | 3 | 10 | 4 | 3 | 4 | 11 | 4 | 4 | 4 | 12 |
| 34 | 3 | 3 | 3 | 9 | 3 | 4 | 3 | 10 | 4 | 3 | 3 | 10 | 4 | 3 | 4 | 11 |
| 35 | 2 | 3 | 3 | 8 | 3 | 3 | 3 | 9 | 3 | 3 | 4 | 10 | 4 | 3 | 3 | 10 |
| 36 | 4 | 3 | 4 | 11 | 4 | 4 | 3 | 11 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| 37 | 3 | 4 | 4 | 11 | 3 | 3 | 3 | 9 | 4 | 4 | 3 | 11 | 3 | 3 | 3 | 9 |
| 38 | 4 | 3 | 3 | 10 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 | 3 | 3 | 3 | 9 |
| 39 | 4 | 4 | 4 | 12 | 3 | 4 | 4 | 11 | 3 | 3 | 3 | 9 | 3 | 3 | 2 | 8 |
| \bar{X} | 3.03 | 3.13 | 3.077 | 9.23 | 2.82 | 2.95 | 2.85 | 8.615 | 3.00 | 3.05 | 2.92 | 8.97 | 3.21 | 3.154 | 3.21 | 9.564 |
| S.D. | 0.84 | 0.77 | 0.807 | 1.83 | 0.60 | 0.72 | 0.78 | 1.771 | 0.46 | 0.56 | 0.53 | 1.14 | 0.70 | 0.709 | 0.77 | 1.535 |

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ง
รายนามผู้ทรงคุณวุฒิที่ตรวจสอบเกณฑ์การให้คะแนน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายนามผู้ทรงคุณวุฒิที่ตรวจสอบเกณฑ์การให้คะแนน

1. รองศาสตราจารย์ อัจฉรา ชีวพันธ์ อาจารย์วิชาภาษาไทย
โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายประถม
2. ผู้ช่วยศาสตราจารย์ สุปราณี ดาราฉาย อาจารย์วิชาภาษาไทย
โรงเรียนสาธิตจุฬาลงกรณ์มหาวิทยาลัย ฝ่ายมัธยม
3. อาจารย์กฤษณีรา บุญเลี้ยง ครูภาษาไทยดีเด่น
โรงเรียนราชวินิต มัธยม
4. อาจารย์ทัศนีย์ หัวหาญ ครูภาษาไทยดีเด่น
โรงเรียนพระฤทัยคอนแวนต์
5. อาจารย์สินีมาศ ดวงสัมพันธ์ ครูภาษาไทยดีเด่น
โรงเรียนชุมชนรัชดาภิเษกมิตรภาพที่ 115



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก จ
รายนามอาจารย์ที่ตรวจให้คะแนนผลงาน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รายนามอาจารย์ที่ตรวจให้คะแนนผลงาน

- | | | |
|--------------------|-------------|----------------------------|
| 1. อาจารย์ศิริวรรณ | ทองทวีวัฒน์ | โรงเรียนไพฑูริศึกษา |
| 2. อาจารย์จวีรัตน์ | กุดระแสง | โรงเรียนไพฑูริศึกษา |
| 3. อาจารย์นพดล | สังข์ทอง | โรงเรียนไพฑูริศึกษา |
| 4. อาจารย์สุภาวดี | คำฝักฝน | โรงเรียนอัสสัมชัญแผนกประถม |
| 5. อาจารย์จินตจุฑา | คำทองสุข | โรงเรียนอัสสัมชัญแผนกประถม |
| 6. อาจารย์อุไรพร | นาคะเสถียร | โรงเรียนอัสสัมชัญแผนกประถม |
| 7. อาจารย์ธนะวดี | ใจบุญ | โรงเรียนศิริวิทยา |
| 8. อาจารย์นิตยา | ช่อนกลิ่น | โรงเรียนศิริวิทยา |



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวปวีณา ปีอาทิตย์ เกิดเมื่อวันที่ 9 มีนาคม พ.ศ. 2520 อยู่บ้านเลขที่ 32/10 ถนนราษฎร์อำนวย ต.ในเวียง อ.เมือง จ. น่าน สำเร็จการศึกษาครุศาสตรบัณฑิต วิชาเอก วิทยาศาสตร์ทั่วไป – เคมี ภาควิชามัธยมศึกษา จากคณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2542 และเข้าศึกษาต่อในหลักสูตร ครุศาสตรมหาบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2543



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย