

การพัฒนากฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

ปริญญานิพนธ์
ของ
ชาญวิทย์ จรัสสุทธิอิศร

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา

เมษายน 2545

ลิขสิทธิ์เป็นของ มหาวิทยาลัยศรีนครินทรวิโรฒ

การพัฒนากฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

บทคัดย่อ
ของ
ชาญวิทย์ จรัสสุทธิอิสร์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
เมษายน 2545

Signature

ชาญวิทย์ จรัสสุทธิอิสร. (2545), การพัฒนากฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์
ขั้นบูรณาการ. ปรินญาณิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ. คณะกรรมการควบคุม: รองศาสตราจารย์ ดร. บุญเชิด
ภิญโญนนตพงษ์, รองศาสตราจารย์นิภา ศรีไพโรจน์

การวิจัยในครั้งนี้ มีจุดมุ่งหมายสำคัญ เพื่อพัฒนากฎเกณฑ์การให้คะแนน (Scoring Rubric) ทักษะ
กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่เชื่อถือได้ และศึกษาการตรวจให้คะแนนสองรูปแบบ คือ รูปแบบ
ที่ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบทุกคน และรูปแบบที่ผู้ตรวจทุกคนตรวจแบบทดสอบ
ทุกทักษะเฉพาะผู้เข้าสอบบางคน

กลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนบุญวัฒนา 2 สังกัดกรมสามัญศึกษา
จังหวัดนครราชสีมา จำนวน 50 คน ผู้ตรวจให้คะแนนแบบทดสอบเป็นครูผู้สอนวิชาวิทยาศาสตร์ในระดับชั้น
มัธยมศึกษาตอนปลาย จำนวน 5 คน เครื่องมือที่ใช้เป็นแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์
ขั้นบูรณาการ จำนวน 5 ทักษะ ผลการวิจัยพบว่า

1. กฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ที่ผู้วิจัยสร้างขึ้น มี
คุณภาพและประสิทธิภาพ โดยการพิจารณาความสอดคล้องของการตรวจให้คะแนนของผู้ตรวจ เมื่อคำนวณ
ค่าความเชื่อมั่นจากผู้ตรวจแต่ละคนโดยใช้สูตรของ Feldt – Raju แล้วปรับขยายโดยใช้สูตรของ Spearman –
Brown เท่ากับ .9677 ความคลาดเคลื่อนมาตรฐานของการวัด เท่ากับ .3957

2. ค่าสัมประสิทธิ์การสรุปอ้างอิง รูปแบบที่ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบ
ทุกคน เท่ากับ .7052 และ รูปแบบที่ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะเฉพาะผู้เข้าสอบบางคน เท่ากับ
.6432 ทดสอบความแตกต่างของการตรวจให้คะแนนทั้งสองรูปแบบ พบว่า แตกต่างกันอย่างมีนัยสำคัญทาง
สถิติที่ .01

DEVELOPMENT OF SCORING RUBRIC FOR INTEGRATED SCIENTIFIC PROCESS SKILLS

AN ABSTRACT

BY

CHARNWIT JARUSSOOTTI-ISORN

Presented in partial fulfillment of the requirements
for the Master of Education degree in Educational Measurement
at Srinakharinwirot University
April 2002

Charnwit Jarussootti-Isorn. (2002). *Development of Scoring Rubric Integrated Scientific Process Skills*. Master Thesis, M.Ed. (Educational Measurement). Bangkok : Graduate School, Srinakharinwirot University. Advisor Committee: Assoc. Prof. Dr. Boonchird Pinyoanuntapong , Assoc. Prof. Nipa Sripairot.

The major purpose of this study is to develop scoring rubric of integrated scientific process skills and to study the reliability of two rating techniques as follows : 1) All raters checked every skill and item of the integrated scientific process test paper of each student. 2) All raters checked every skill and item of integrated science process test of some students.

The sampling of this study is fifty Mattayom Suksa four students at Boonwattata 2 School, General Education Department, Nakhonratchasima Province and five science teachers who teach in senior high school level. The instrument used in this study is an integrated scientific skills test which consisted of five scientific skills. The study results are as follows:

1. The scoring rubric created by the researcher was used in order to make the scoring result of each rater balanced . When analyzing reliability of the test which referred the result to scoring techniques ,the reliability of each rater used the Feldt – Raju formula and increased by Spearman – Brown formula was at .9677. The Standard error of measurement was .3957

2. The reliability of scoring techniques calculated by using referenced theory and checking by all raters was at .7052. The technique used for checking some students by all raters was at .6432. The differences of scoring was statistically significant at .01.

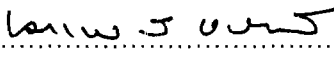
ปริญญานิพนธ์นี้ ได้รับทุนอุดหนุนการทำวิจัยจาก
สาขาวิจัย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี
ประจำปี 2545

ปริญญานิพนธ์
เรื่อง

การพัฒนากฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

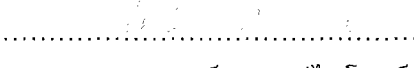
ของ
นายชาญวิทย์ จรัสสุทธิอิศร

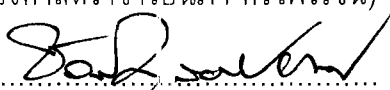
ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาการศึกษามหาบัณฑิต วิชาเอกการวัดผลการศึกษา
ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.......... คณบดีบัณฑิตวิทยาลัย
(รองศาสตราจารย์ ดร. นภาพร หะวานนท์)
วันที่ 30 เมษายน พ.ศ. 2545

คณะกรรมการสอบปริญญานิพนธ์

.......... ประธาน
(รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนตพงษ์)

.......... กรรมการ
(รองศาสตราจารย์ นipa ศรีไพโรจน์)

.......... กรรมการที่แต่งตั้งเพิ่มเติม
(อาจารย์ชวลิต รวยอาจิด)

.......... กรรมการที่แต่งตั้งเพิ่มเติม
(ดร. สุพร เข้มเฮง)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จสมบูรณ์ได้ เพราะได้รับความช่วยเหลือจาก รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนท์พงษ์ ประธานคณะกรรมการควบคุม และรองศาสตราจารย์ นิภา ศรีไพโรจน์ กรรมการ ที่ได้ให้คำปรึกษา ชี้แนะและตรวจสอบ จนสามารถบรรลุจุดหมาย ผู้วิจัยรู้สึกซาบซึ้งในความอนุเคราะห์ และเสียสละเวลาเป็นอย่างมากและขอกราบขอบพระคุณท่านเป็นอย่างสูง

ขอกราบขอบพระคุณ ดร. สุพร เข้มแข็ง และ อาจารย์ชวลิต รวยอาจิณ อาจารย์ประจำภาควิชา การวัดผลและวิจัยการศึกษา กรรมการที่แต่งตั้งเพิ่มเติม ที่ได้ให้คำแนะนำ ชี้แนะ ในการปรับปรุงรูปแบบและ เนื้อหาสาระที่ถูกต้องชัดเจนยิ่งขึ้น เพื่อให้งานวิจัยนี้มีคุณภาพที่ดี ผู้วิจัยขอกราบขอบพระคุณท่านเป็นอย่าง สูง

ขอกราบขอบพระคุณ คณาจารย์ภาควิชาการวัดผลและวิจัยการศึกษา มหาวิทยาลัย ศรีนครินทรวิโรฒ ทุกท่าน ที่ได้ช่วยเหลือ ดูแล ให้คำปรึกษาแนะแนวทางในการทำวิจัยตลอดมา ผู้วิจัยขอ กราบขอบพระคุณทุกท่านเป็นอย่างสูง

ขอกราบขอบพระคุณ ดร. สุพร เข้มแข็ง และ อาจารย์ชวลิต รวยอาจิณ อาจารย์ประจำภาควิชา การวัดผลและวิจัยการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ผู้ช่วยศาสตราจารย์กมลทิพย์ กลสิมาร์ อาจารย์ ประจำภาควิชาชีววิทยา สถาบันราชภัฏนครราชสีมา ดร.วตาภรณ์ พูนผลอำนวย ฝ่ายนิเทศงานวัดผลและ ประเมินผล และ อาจารย์ประเสริฐ ทิศกลาง หัวหน้าหน่วยศึกษานิเทศก์ กรมสามัญศึกษา เขตการศึกษา 11 ที่ได้ให้ความอนุเคราะห์ตรวจสอบเครื่องมือที่ใช้ในงานวิจัย

ขอขอบคุณ ผู้บริหาร ครู-อาจารย์ และนักเรียนโรงเรียนบุญวัฒนา 2 จังหวัดนครราชสีมา ที่ได้ให้ ความร่วมมือในการเก็บรวบรวมข้อมูลในการวิจัย

ขอขอบคุณ หน่วยศึกษานิเทศก์ กรมสามัญศึกษา เขตการศึกษา 11 เพื่อนร่วมงานและครอบครัว ที่ให้กำลังใจ และช่วยเหลือเป็นอย่างดีตลอดระยะเวลาที่ทำการศึกษาและวิจัย ขอขอบคุณเพื่อน ๆ นิสิต ปริญญาโท สาขาการวัดผลการศึกษา ภาคพิเศษ รุ่นที่ 11 ทุกคนที่ให้อาการใจ ร่วมมือค้นคว้าข้อมูลด้วยกัน ตลอดมา

ขอขอบคุณ สาขาวิจัย สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี ที่มองเห็นความ สำคัญและประโยชน์ที่ครูผู้สอนวิทยาศาสตร์จะได้รับจากงานวิจัยในครั้งนี้ ได้อนุเคราะห์ทุนอุดหนุนการวิจัย ประจำปีงบประมาณ 2545 มาสนับสนุนเป็นทุนดำเนินการ ผู้วิจัยรู้สึกซาบซึ้งและภาคภูมิใจที่ได้รับความ อนุเคราะห์ในครั้งนี้ และขอกราบขอบพระคุณเป็นอย่างสูง

คุณค่าและประโยชน์ใด ๆ ที่ได้รับจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบให้เป็นเครื่องบูชา พระคุณบิดา มารดา ครู-อาจารย์ และผู้มีพระคุณทุกท่าน

ชาญวิทย์ จรัสสุทธิอิสร์

สารบัญ

บทที่

หน้า

1 บทนำ	1
ภูมิหลัง	1
ความมุ่งหมายของการวิจัย	4
ความสำคัญของการวิจัย	4
ขอบเขตของการวิจัย	4
ประชากรและกลุ่มตัวอย่างที่ใช้ในการวิจัย	4
รูปแบบการทดลองที่ใช้ในการวิจัย	4
ตัวแปรที่ใช้ในการวิจัย	5
นิยามศัพท์เฉพาะ	5
 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	8
ทักษะกระบวนการทางวิทยาศาสตร์	8
ความหมายของทักษะกระบวนการทางวิทยาศาสตร์	8
ประเภทของทักษะกระบวนการทางวิทยาศาสตร์	9
ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ	17
การประเมินผลทักษะกระบวนการทางวิทยาศาสตร์	30
แบบทดสอบความเรียง	31
ความหมายของแบบทดสอบความเรียง	31
การสร้างแบบทดสอบความเรียง	33
องค์ประกอบที่ส่งผลต่อคะแนนในแบบทดสอบความเรียง	37
กฎเกณฑ์การให้คะแนน	38
ความหมายของกฎเกณฑ์การให้คะแนน	38
ประเภทของกฎเกณฑ์การให้คะแนน	39
การสร้างกฎเกณฑ์การให้คะแนน	41
แนวทางในการตรวจให้คะแนน	46
ทฤษฎีค่าความเชื่อมั่น	50
การประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม	50
การประมาณค่าความเชื่อมั่นที่เกี่ยวกับการตรวจให้คะแนน	52
การประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง	55
งานวิจัยที่เกี่ยวข้อง	63
สมมติฐานในการวิจัย	78

สารบัญ(ต่อ)

บทที่

หน้า

3 วิธีดำเนินการวิจัย.....	79
การกำหนดประชากรและเลือกกลุ่มตัวอย่าง	79
รูปแบบการทดลองที่ใช้ในการวิจัย	80
เครื่องมือที่ใช้ในการวิจัย.....	80
วิธีดำเนินการเก็บรวบรวมข้อมูล	85
การจัดกระทำและการวิเคราะห์ข้อมูล	85
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	86
 4 ผลการวิเคราะห์ข้อมูล.....	 89
 5 สรุปผล อภิปราย และข้อเสนอแนะ.....	 96
ความมุ่งหมายของการวิจัย	96
สมมติฐานของการวิจัย	96
ประชากรที่ใช้ในการวิจัย	96
กลุ่มตัวอย่างที่ใช้ในการวิจัย	96
เครื่องมือที่ใช้ในการวิจัย.....	96
วิธีดำเนินการเก็บรวบรวมข้อมูล	97
สรุปผลการวิเคราะห์ข้อมูล.....	98
อภิปรายผล.....	99
ข้อเสนอแนะ	100
 บรรณานุกรม	 101
 ภาคผนวก.....	 111
 ประวัติย่อผู้วิจัย	 138

บัญชีตาราง

ตาราง	หน้า
1 ผลการทดลองละลายสาร A ในน้ำ 50 cm ³ ที่อุณหภูมิต่างกัน.....	19
2 ผลการทดลองกิจกรรมชุดกลิ้งหูกะเมน	29
3 ผลการวัดความดันบรรยากาศและความหนาแน่นของอากาศที่ระดับความสูงต่าง ๆ	29
4 เกณฑ์การให้คะแนนแบบรวมองค์ประกอบ.....	42
5 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ	43
6 คำอธิบายระดับคุณภาพ "การพูดอภิปราย".....	44
7 การกำหนดคะแนนการตัดสินระดับคุณภาพ	45
8 ผลการวิเคราะห์ความแปรปรวนจากการตรวจให้คะแนนแบบสมบูรณ์	54
9 ผลการวิเคราะห์ความแปรปรวนจากการตรวจให้คะแนนแบบไม่สมบูรณ์.....	54
10 การประมาณค่าความแปรปรวนขึ้นมาใหม่ของแบบแผน $p \times i \times R$	62
11 ผลการวิเคราะห์พฤติกรรมที่บ่งชี้ทักษะกระบวนการทางวิทยาศาสตร์.....	82
12 ค่าสถิติพื้นฐานของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ	90
13 ค่าความเชื่อมั่นของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ	91
14 ค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของ แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ เมื่อความสัมพันธ์ ขององค์ประกอบเป็น $i \times r \times p$	92
15 ค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของ แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ เมื่อความสัมพันธ์ ขององค์ประกอบเป็น $i \times (p : r)$	93
16 ค่าความแปรปรวนของคะแนนเอกภพ คะแนนความคลาดเคลื่อน และสัมประสิทธิ์การสรุป อ้างอิงของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ตรวจ ให้คะแนนโดยใช้แบบแผน $i \times r \times p$	94
17 ค่าความแปรปรวนของคะแนนเอกภพ คะแนนความคลาดเคลื่อน และสัมประสิทธิ์การสรุป อ้างอิงของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ตรวจ ให้คะแนนโดยใช้แบบแผน $i \times (p : r)$	94
18 การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของการตรวจให้คะแนนแบบทดสอบวัด ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการทั้งสองรูปแบบ.....	95
19 ดัชนีความสอดคล้องของการประเมินจากผู้เชี่ยวชาญในการตรวจสอบคุณภาพของ เครื่องมือ (IOC)	114
20 ค่าความยาก (p) ค่าอำนาจจำแนก (r_{xy}) ของแบบทดสอบวัดทักษะกระบวนการทาง วิทยาศาสตร์ขั้นบูรณาการ	115
21 ผลการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ตามแบบแผน $i \times r \times p$	131

บัญชีตาราง(ต่อ)

ตาราง	หน้า
22 ผลการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ตามแบบแผน $i \times (p : r)$	136

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 แผนภูมิ Venn แสดงองค์ประกอบความแปรปรวนของแบบแผน $p \times i \times r$ เมื่อองค์ประกอบที่ศึกษาทั้งหมดเป็นองค์ประกอบสุ่ม.....	60
2 แผนภูมิ Venn แสดงองค์ประกอบความแปรปรวนของแบบแผน $p \times i \times r$ เมื่อ p และ i เป็นองค์ประกอบสุ่ม และ r เป็นองค์ประกอบกำหนด.....	61
3 แสดงลำดับในการสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์.....	81

บทที่ 1

บทนำ

ภูมิหลัง

ปัจจุบันวิทยาศาสตร์และเทคโนโลยีมีความสำคัญเกี่ยวข้องกับชีวิตของมนุษย์มาก วิทยาศาสตร์ช่วยให้เราได้รับความสะดวกสบาย ความก้าวหน้าทางวิทยาศาสตร์ดำเนินไปอย่างรวดเร็ว จนทำให้วิถีชีวิตของมนุษย์เกือบทุกมุมโลกต้องเกี่ยวข้องกับวิทยาศาสตร์และเทคโนโลยีอย่างไม่มีทางหลีกเลี่ยง วิทยาศาสตร์ไม่ได้หมายถึงความรู้เพียงอย่างเดียว แต่ยังรวมถึงวิธีการแสวงหาความรู้ที่มีระบบ ความรู้และวิธีการแสวงหาความรู้ทางวิทยาศาสตร์นับเป็นส่วนประกอบที่สำคัญในการพัฒนาประเทศชาติ ฉะนั้น การพัฒนาการเรียนการสอนวิทยาศาสตร์ให้เกิดประสิทธิภาพในโรงเรียน ถือว่าเป็นรากฐานที่สำคัญ (ทบวงมหาวิทยาลัย. 2525 : 58) วิทยาศาสตร์มีความสำคัญในการพัฒนาความสามารถในด้านการใช้เหตุผล การตัดสินใจ ตลอดจนมีเจตคติเชิงวิทยาศาสตร์ ได้บุคคลที่มีคุณภาพสามารถพัฒนาสังคมและประเทศชาติ ให้เจริญรุ่งเรืองได้ตลอดเวลา ทุกประเทศในโลกจึงได้จัดหลักสูตรวิทยาศาสตร์ไว้ในโรงเรียนทุกระดับ โดยเฉพาะระดับมัธยมศึกษาชั้นวิทยาศาสตร์จัดเป็นพื้นฐานที่ทุกคนต้องได้เรียน และเป็นที่ประจักษ์ชัดว่า ประเทศที่มีความมั่นคง สามารถพัฒนาเศรษฐกิจได้ดั้นนั้น เนื่องจากประชากรมีความรู้ความสามารถในเรื่องวิทยาศาสตร์และเทคโนโลยี

การที่จะพัฒนาวิทยาศาสตร์ให้มีประสิทธิภาพนั้น องค์ประกอบที่สำคัญมากอย่างหนึ่ง คือ ครูผู้สอนวิทยาศาสตร์จะต้องมีความรู้ความเข้าใจในความหมายและขอบเขตของวิทยาศาสตร์ ตระหนักถึงความสำคัญของกระบวนการแสวงหาความรู้ทางวิทยาศาสตร์ การค้นคว้าทดลอง เพื่อค้นหาข้อเท็จจริง หลักการ กฎ และทฤษฎี โดยเฉพาะขณะที่ทำการค้นคว้าทดลอง ผู้ทดลองมีโอกาสดูฝึกฝนทั้งในด้านการปฏิบัติและพัฒนาความคิดควบคู่กันไป เช่น ฝึกการสังเกต บันทึกข้อมูล หาความสัมพันธ์ของตัวแปร ตั้งสมมุติฐานและทำการทดลอง เป็นต้น พฤติกรรมที่เกิดขึ้นจากปฏิบัติการและการฝึกฝนความนึกคิดอย่างเป็นระบบนี้ เรียกว่า "ทักษะกระบวนการทางวิทยาศาสตร์" (ทบวงมหาวิทยาลัย. 2525 : 59)

กรมวิชาการ กระทรวงศึกษาธิการ (2533 : 33) ได้กำหนดจุดประสงค์ของหลักสูตรวิทยาศาสตร์ระดับมัธยมศึกษาตอนปลายไว้ดังนี้

1. เพื่อให้มีความเข้าใจในหลักการ ทฤษฎีที่เป็นพื้นฐานของวิชาวิทยาศาสตร์
2. เพื่อให้มีความเข้าใจในลักษณะ ขอบเขตและข้อจำกัดของวิทยาศาสตร์
3. เพื่อให้มีทักษะที่สำคัญในการศึกษาค้นคว้าและคิดค้นทางวิทยาศาสตร์และเทคโนโลยี
4. เพื่อให้มีเจตคติทางวิทยาศาสตร์
5. เพื่อให้ตระหนักถึงความสัมพันธ์ระหว่างวิทยาศาสตร์ เทคโนโลยี มวลมนุษย์และ

สภาพแวดล้อม ในเชิงที่มีอิทธิพลและผลกระทบซึ่งกันและกัน

6. เพื่อนำความรู้ความเข้าใจในเรื่องวิทยาศาสตร์และเทคโนโลยี ไปใช้ให้เกิดประโยชน์ต่อสังคมและการดำรงชีวิตอย่างมีคุณค่า

จุดประสงค์ของหลักสูตรวิทยาศาสตร์ ระดับมัธยมศึกษาตอนปลายดังกล่าว เน้นให้เกิดทักษะกระบวนการทางวิทยาศาสตร์ทั้ง 13 ทักษะ ได้แก่

1. ทักษะการสังเกต
2. ทักษะการจำแนกประเภท
3. ทักษะการวัด
4. ทักษะการคำนวณ
5. ทักษะการหาความสัมพันธ์ระหว่างสเปกกับสเปสและสเปสกับเวลา
6. ทักษะการจัดกระทำข้อมูลเพื่อสื่อความหมาย
7. ทักษะการลงความคิดเห็นจากข้อมูล
8. ทักษะการพยากรณ์
9. ทักษะการตั้งสมมุติฐาน
10. ทักษะการกำหนดและควบคุมตัวแปร
11. ทักษะการกำหนดนิยามเชิงปฏิบัติการ
12. ทักษะการตีความหมายข้อมูลและลงข้อสรุป
13. ทักษะการทดลอง

มีผลการศึกษาที่เกี่ยวข้องกับทักษะกระบวนการทางวิทยาศาสตร์หลายคน เช่น มาโนช วาตะพุกณะ (2523 : 70) ได้ศึกษาผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์และพฤติกรรมด้านทักษะปฏิบัติของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 2 โรงเรียนบุญวัฒนา จังหวัดนครราชสีมา พบว่า ผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์ต่ำกว่าเกณฑ์ที่ควรจะเป็นหรือคิดเป็นร้อยละ 44.28 กิ่งฟ้า สินธุวงษ์ และคณะ (2525 : 113) ศึกษาทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษา พบว่า ผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 และ 4 ของภาคตะวันออกเฉียงเหนือส่วนใหญ่ยังต่ำกว่าเกณฑ์ ชีระชัย นนทพิภักดิ์ (2530 : 117) ศึกษาองค์ประกอบบางประการที่มีอิทธิพลต่อทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จังหวัดอุดรธานี พบว่า คะแนนเฉลี่ยผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์คิดเป็นร้อยละ 41.68 ต่ำกว่าร้อยละ 50 ของคะแนนเต็ม สอดคล้องกับ ปรีชา เจริญ (2531 : 63-64) เปรียบเทียบทักษะกระบวนการทางวิทยาศาสตร์และเจตคติเชิงวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 พบว่า คะแนนเฉลี่ยด้านทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนต่ำกว่าร้อยละ 50 ของคะแนนเต็ม นักเรียนไม่ถึงร้อยละ 50 ได้คะแนนเท่ากับหรือสูงกว่าร้อยละ 50 ของคะแนนเต็ม

จากการศึกษาขั้นต้นพบว่า ความสามารถการใช้ทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนยังไม่บรรลุผลเท่าที่ควร ผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานของนักเรียนระดับชั้นมัธยมศึกษาตอนต้นส่วนมากยังต่ำกว่าร้อยละ 50 ของคะแนนเต็ม ชูติญา สุริยมณฑล (2535 : 68, 74-75) สำรวจประเภทของทักษะกระบวนการทางวิทยาศาสตร์ในหนังสือเรียนระดับมัธยมศึกษาตอนต้น จำนวน 6 เล่ม โดยใช้ ทักษะกระบวนการทางวิทยาศาสตร์ของสมาคมอเมริกันเพื่อความก้าวหน้าในทางวิทยาศาสตร์ (AAAS) 13 ทักษะเป็นเกณฑ์ พบว่า ทักษะการลงความคิดเห็นจากข้อมูลมากที่สุด และทักษะการตั้งสมมุติฐานน้อยที่สุด หนังสือเรียนวิทยาศาสตร์ระดับมัธยมศึกษาตอนต้นทุกเล่มทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานคิดเป็นร้อยละ 78.55 และทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการคิดเป็นร้อยละ 21.45

ผลการสำรวจดังกล่าว ชี้ให้เห็นว่า ทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาในปัจจุบัน นอกจากจะยังไม่บรรลุเป้าหมายตามที่ต้องการแล้ว ยังมีทักษะกระบวนการทางวิทยาศาสตร์ที่ต้องได้รับการพัฒนาเป็นการเร่งด่วน คือ ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ เพราะว่า มีสัดส่วนอยู่ในหนังสือเรียนวิชาวิทยาศาสตร์น้อยมาก เมื่อเปรียบเทียบกับทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน

การแก้ปัญหาดังกล่าวข้างต้นได้ วิธีการหนึ่ง คือการพัฒนาวิธีการประเมินผลให้มีคุณภาพ แต่การวัดผลและประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ก็ยังมีปัญหา ในทางปฏิบัติ ยังไม่เป็นรูปธรรมที่สามารถตรวจสอบได้ ผู้สอนส่วนใหญ่ขาดความรู้ความเข้าใจเกี่ยวกับการประเมินผลด้านทักษะกระบวนการทางวิทยาศาสตร์ (สุวัฒน์ นิยมคำ, 2531) ขาดความเข้าใจในการเลือกวิธีการวัดผลและประเมินผลและวิธีสร้างกฎเกณฑ์การให้คะแนน (Scoring Rubric) ที่เป็นปรนัย ให้มีความเชื่อมั่นสูงและลดความคลาดเคลื่อนจากการวัดในแหล่งต่าง ๆ ให้เหลือน้อยที่สุด โดยมากครูผู้สอนจะนิยมวัดผลสัมฤทธิ์ด้านความรู้ความจำ โดยใช้แบบทดสอบที่มีตัวเลือกกำหนดให้ตอบ ที่สะดวกทั้งวิธีการตรวจให้คะแนนและสามารถนำผลไปวิเคราะห์หาคุณภาพของแบบทดสอบ แต่วิธีการดังกล่าวนี้ เราอาจไม่ทราบถึงพัฒนาการที่แท้จริงเกี่ยวกับวิธีการแสวงหาความรู้ของนักเรียนและทักษะพื้นฐานความรู้ทางวิทยาศาสตร์ที่จำเป็น ส่งผลให้ผลสัมฤทธิ์ทางการเรียนในวิชาวิทยาศาสตร์ยังคงตกต่ำเหมือนกับที่เคยผ่านมา การพัฒนากฎเกณฑ์การตรวจให้คะแนน เพื่อใช้เป็นแนวทางให้ครูผู้สอนได้ใช้ประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ และพัฒนาศักยภาพในการเรียนรู้วิธีการแสวงหาความรู้ของผู้เรียน เพื่อให้ทันต่อการเปลี่ยนแปลงของหลักสูตรตามพระราชบัญญัติการศึกษา พ.ศ. 2542 จึงนับเป็นสิ่งที่ต้องกระทำเป็นอันดับแรก

ตัวประกอบที่มีผลต่อค่าความเชื่อมั่นของแบบประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ ที่มีลักษณะเป็นคำถามปลายเปิดประกอบด้วย ผู้ตรวจให้คะแนน วิธีการตรวจ จำนวนแบบทดสอบ จำนวนผู้ตอบ และทฤษฎีที่ประมาณค่าความเชื่อมั่น โดยเฉพาะความไม่แน่นอนของผู้ตรวจให้คะแนน (Rater Unreliability) นับเป็นปัญหาที่สำคัญ ความไม่แน่นอนของผู้ตรวจให้คะแนนจะส่งผลให้คะแนนแปรเปลี่ยนไปตามผู้ตรวจให้คะแนนแต่ละคน แม้แต่ผู้ตรวจให้คะแนนที่เป็นคนเดียวกัน ก็จะทำให้คะแนนแปรเปลี่ยนไปตามระยะเวลาที่ทำการตรวจ (บุญเชิด ภิญโญอนันตพงษ์, 2533) ทฤษฎีที่ใช้ในการประเมินค่าความเชื่อมั่นเป็นอีกปัญหาหนึ่งที่มีความสำคัญ ทฤษฎีการวัดแบบดั้งเดิม (Classical Test Theory : CTT) จะประมาณค่าความเชื่อมั่นโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนของผู้ตอบภายใต้เงื่อนไขการวัดที่เท่าเทียม ได้แก่ ค่าเฉลี่ยเท่ากัน ความแปรปรวนที่เท่ากัน และค่าสหสัมพันธ์ระหว่างแบบทดสอบเท่ากัน ซึ่งในทางปฏิบัติ การวัดผลและประเมินผลโดยทั่วไป ไม่สามารถจัดหาเครื่องมือที่มีคุณสมบัติของความเท่าเทียมอย่างสมบูรณ์ได้ โดยเฉพาะถ้าต้องคำนึงผู้ตรวจให้คะแนนด้วยแล้ว การใช้ทฤษฎีการวัดแบบดั้งเดิมจึงไม่เหมาะสม ด้วยข้อจำกัดดังกล่าว ครอนบัคและคณะ (Cronbach, et al. 1972) ได้พัฒนาทฤษฎีการสรุปอ้างอิง (Generalizability Theory) ขึ้นมา โดยทฤษฎีนี้จะไม่ยึดข้อตกลงเกี่ยวกับคุณสมบัติที่เท่าเทียม ทฤษฎีนี้ได้ปรับปรุงพัฒนาจากทฤษฎีการวัดแบบดั้งเดิม ใช้วิธีการวิเคราะห์ความแปรปรวนหาความคลาดเคลื่อนจากหลาย ๆ แหล่ง (Brennan, 1983 : 1)

จากเหตุผลดังกล่าว ผู้วิจัยจึงได้นำมาใช้เป็นแนวทางการพัฒนา การประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ โดยการเลือกตัวแปรข้างต้นมาเป็นประเด็นในการวิจัย อาศัยความเห็นที่สอดคล้องกันของผู้ตรวจและการลดความคลาดเคลื่อนในการวัดให้เหลือน้อยที่สุด เป็นตัวบ่งชี้ความน่าเชื่อถือของกฎเกณฑ์การให้คะแนน (Scoring Rubric) พร้อมทั้งศึกษาการลดภาระของการตรวจให้คะแนนโดยใช้ค่าสัมประสิทธิ์การสรุปอ้างอิง เพื่อใช้เป็นแนวทางในการสร้างกฎเกณฑ์การให้คะแนน (Scoring Rubric) ในด้านอื่น ๆ อีกต่อไป

ความมุ่งหมายของการวิจัย

1. เพื่อพัฒนากฎเกณฑ์การให้คะแนน (Scoring Rubric) ทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
2. เพื่อศึกษารูปแบบของการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการสองรูปแบบที่คำนวณโดยใช้ทฤษฎีการสุ่มอย่างอิง

ความสำคัญของการวิจัย

การวิจัยครั้งนี้ เป็นแนวทางการพัฒนาและหาความเชื่อมั่นของกฎเกณฑ์การให้คะแนนด้านทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ จะเป็นทางเลือกหนึ่งในการหาประสิทธิภาพของกฎเกณฑ์การให้คะแนนตามแนวทางการจัดประสบการณ์การเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ ยังช่วยส่งเสริมให้มีการพัฒนาการวัดผลและประเมินผลในด้านอื่น ๆ ให้แพร่หลายต่อไป

ขอบเขตของการวิจัย

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนมัธยมศึกษาสังกัดกรมสามัญศึกษาจังหวัดนครราชสีมา ปีการศึกษา 2544 จำนวน 16,919 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้า เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ปีการศึกษา 2544 โรงเรียนบุญวัฒนา 2 สังกัดสำนักงานสามัญศึกษาจังหวัดนครราชสีมา จำนวน 50 คน ซึ่งเลือกมาโดยการสุ่มแบบหลายขั้นตอน (Multi - Stage Random Sampling)

รูปแบบการทดลองที่ใช้ในการวิจัย

การศึกษาค้นคว้าในครั้งนี้ ผู้วิจัยกำหนดความสัมพันธ์ขององค์ประกอบที่ศึกษา ในชั้นการศึกษาเพื่อการสุ่มอย่างอิง เป็น 2 รูปแบบ คือ

1. แบบแผน $i \times r \times p$ เป็นแบบแผนการวิเคราะห์ค่าความเชื่อมั่นของคะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ได้จาก ผู้เข้าสอบทุกคนสอบด้วยแบบทดสอบชุดเดียวกัน และผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบทุกคน
2. แบบแผน $i \times (p : r)$ เป็นแบบแผนการวิเคราะห์ค่าความเชื่อมั่นของคะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ได้จาก ผู้เข้าสอบทุกคนสอบด้วยแบบทดสอบชุดเดียวกัน และผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบบางคน แต่แยกตรวจผู้เข้าสอบเป็นรายกลุ่มโดยไม่ซ้ำกัน

ตัวแปรที่ใช้ในการวิจัย

1. ตัวแปรอิสระ ได้แก่ รูปแบบการตรวจให้คะแนน

1.1 ผู้ตรวจทุกคนตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ทุกทักษะของผู้เข้าสอบทุกคน

1.2 ผู้ตรวจทุกคนตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ทุกทักษะของผู้เข้าสอบบางคน โดยไม่ซ้ำกัน

2. ตัวแปรตาม ได้แก่

ค่าความเชื่อมั่นของคะแนนทักษะกระบวนการทางวิทยาศาสตร์

นิยามศัพท์เฉพาะ

1. ทักษะกระบวนการทางวิทยาศาสตร์ (Science Process Skills) หมายถึง พฤติกรรมที่เกิดจากการปฏิบัติและฝึกฝนความนึกคิดอย่างเป็นระบบ เป็นความสามารถในการ แสวงหาความรู้ทางวิทยาศาสตร์ แบ่งออกเป็น 2 ประเภท คือ ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน (Basic Science Process Skills) และ ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ (Integrated Science Process Skills)

2. ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน (Basic Science Process Skills) หมายถึง ปฏิบัติการหรือวิธีการแสวงหาความรู้ทางวิทยาศาสตร์ ก่อให้เกิดพัฒนาการทางด้านสติปัญญา ประกอบด้วย 8 ทักษะ ได้แก่ ทักษะการสังเกต ทักษะการจำแนกประเภท ทักษะการวัด ทักษะการคำนวณ ทักษะการหาความสัมพันธ์ระหว่างสเปกกับสเปสและสเปสกับเวลา ทักษะการจัดกระทำข้อมูลและสื่อความหมาย ทักษะการลงความเห็นจากข้อมูล และทักษะการพยากรณ์ มีรายละเอียดดังนี้

2.1 ทักษะการสังเกต (Observing) หมายถึง ความสามารถในการใช้ประสาทสัมผัสอย่างใดอย่างหนึ่งหรือหลายอย่างรวมกัน ซึ่งได้แก่ ตา หู จมูก ลิ้น และกายสัมผัส เข้าสัมผัสโดยตรงกับวัตถุ

2.2 ทักษะการจำแนกประเภท (Classifying) หมายถึง ความสามารถในการจัดแบ่งหรือเรียงลำดับวัตถุหรือสิ่งที่อยู่ในปรากฏการณ์ต่างๆ ออกเป็นพวกๆ โดยมีเกณฑ์ในการจัดแบ่ง

2.3 ทักษะการวัด (Measuring) หมายถึง ความสามารถในการใช้เครื่องมือวัดปริมาณของสิ่งต่าง ๆ ได้อย่างถูกต้อง โดยมีหน่วยกำกับเสมอ รวมไปถึงการเลือกใช้เครื่องมือวัดได้อย่างถูกต้องเหมาะสมต่อสิ่งที่ต้องการวัด

2.4 ทักษะการคำนวณ (Using Numbers) หมายถึง ความสามารถในการบวก ลบ คูณ และหารตัวเลขที่แสดงค่าปริมาณของสิ่งใดสิ่งหนึ่งซึ่งได้จากการสังเกต การวัด การทดลองโดยตรง หรือจากแหล่งอื่น ๆ อีกทอดหนึ่ง

2.5 ทักษะการหาความสัมพันธ์ระหว่างสเปกกับสเปสและสเปสกับเวลา (Space/Space and Space/Time Relationships) หมายถึง ความสามารถในการหาความสัมพันธ์ระหว่าง 3 มิติกับ 2 มิติระหว่างตำแหน่งของวัตถุหนึ่งกับอีกวัตถุหนึ่ง ระหว่างสเปสของวัตถุกับเวลา

2.6 ทักษะการจัดกระทำข้อมูลและสื่อความหมาย (Organizing Data and Communicating) หมายถึง ความสามารถในการนำข้อมูลดิบที่ได้จากการสังเกต การทดลอง หรือจากแหล่งอื่นที่มีข้อมูลดิบอยู่แล้วมาจัดกระทำเสียใหม่ โดยอาศัยวิธีการต่าง ๆ เช่น การจัดเรียงลำดับ การจัดแยกประเภท การหาค่าเฉลี่ย เป็นต้น

2.7 ทักษะการลงความเห็นจากข้อมูล (Inferring) หมายถึง ความสามารถในการนำข้อมูลที่ได้อาจจากการสังเกตวัตถุหรือปรากฏการณ์ไปสัมพันธ์กับความรู้หรือประสบการณ์เดิมเพื่อลงข้อสรุป หรืออธิบายปรากฏการณ์หรือวัตถุนั้น

2.8 ทักษะการพยากรณ์ (predicting) หมายถึง ความสามารถในการทำนายหรือคาดคะเนสิ่งที่จะเกิดขึ้นล่วงหน้า โดยอาศัยการสังเกตปรากฏการณ์ที่เกิดขึ้นซ้ำ ๆ หรือความรู้ที่เป็นหลักการ กฎหรือทฤษฎีในเรื่องนั้นมาช่วยในการทำนาย

3. ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ (Integrated Science Process Skills)

หมายถึง วิธีการแสวงหาความรู้ทางวิทยาศาสตร์ เป็นพฤติกรรมที่เกิดจากการปฏิบัติและฝึกฝนความคิดอย่างเป็นระบบ โดยอาศัยการผสมผสานจากทักษะกระบวนการขั้นพื้นฐาน ประกอบด้วย 5 ทักษะ ได้แก่ ทักษะการตั้งสมมุติฐาน ทักษะการกำหนดและควบคุมตัวแปร ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการตีความหมายข้อมูลและลงข้อสรุป และทักษะการทดลอง มีรายละเอียดดังนี้

3.1 ทักษะการตั้งสมมุติฐาน (Formulating Hypothesis) หมายถึง ความสามารถในการให้คำอธิบายซึ่งเป็นคำตอบล่วงหน้าก่อนที่จะดำเนินการทดลองเพื่อตรวจสอบความถูกต้องเป็นจริงในเรื่องนั้น ๆ โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การพยากรณ์ ประกอบกับประสบการณ์ กฎ หลักการ และทฤษฎีที่เกี่ยวข้อง

3.2 ทักษะการกำหนดและควบคุมตัวแปร (Identifying and Controlling) หมายถึง ความสามารถในการบ่งชี้และกำหนดตัวแปรต้น ตัวแปรตาม และควบคุมตัวแปรที่เป็นสาเหตุอื่น ๆ ในปรากฏการณ์หนึ่ง ๆ ที่ต้องการศึกษา โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การพยากรณ์ และการลงความเห็นจากข้อมูล เป็นต้น

3.3 ทักษะการกำหนดนิยามเชิงปฏิบัติการ (Defining Operationally) หมายถึง ความสามารถในการกำหนดความหมายและขอบเขตของคำหรือตัวแปรต่าง ๆ ในปฏิบัติการทดลองให้สามารถทดสอบหรือวัดได้ โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การวัด เป็นต้น

3.4 ทักษะการตีความหมายข้อมูลและลงข้อสรุป (Interpreting Data and Making Conclusion) หมายถึง ความสามารถในการแปลความหมายหรือบรรยายลักษณะและคุณสมบัติของข้อมูล พร้อมทั้งสรุปความสัมพันธ์ของตัวแปรที่อยู่ในข้อมูลที่ต้องการศึกษา โดยอาศัยทักษะกระบวนการขั้นพื้นฐานที่สำคัญ ได้แก่ การสังเกต การวัด การจัดกระทำและสื่อความหมายข้อมูล และการลงความเห็นจากข้อมูล เป็นต้น

3.5 ทักษะการทดลอง (Experimenting) หมายถึง ความสามารถในการดำเนินการตรวจสอบสมมุติฐานโดยการทดลอง ซึ่งเริ่มตั้งแต่การออกแบบการทดลอง การปฏิบัติการทดลองตามขั้นตอนที่ออกแบบไว้ ตลอดจนการใช้วัสดุอุปกรณ์ได้อย่างถูกต้อง โดยอาศัยทักษะกระบวนการขั้นพื้นฐานและทักษะกระบวนการขั้นบูรณาการทุกทักษะ

4. กฎเกณฑ์การให้คะแนน (Scoring Rubric) หมายถึง แนวทางที่กำหนดขึ้นเพื่อชี้บอกระดับของทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการในแต่ละด้าน ซึ่งข้อมูลได้มาจากการพิจารณาผลของผู้ตรวจให้คะแนนเพื่อให้เกิดความเข้าใจตรงกัน การวิจัยในครั้งนี้ ผู้วิจัยได้พัฒนากฎเกณฑ์การให้คะแนน (Scoring Rubric) โดยใช้รูปแบบเฉพาะ (Specific Rubric) ที่กำหนดคุณภาพเป็น 4 ระดับ

5. รูปแบบการตรวจให้คะแนน หมายถึง ลักษณะของการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ แบ่งออกเป็น 2 รูปแบบ คือ ผู้ตรวจทุกคนตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ทุกทักษะของผู้เข้าสอบทุกคน และผู้ตรวจทุกคนตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ทุกทักษะของผู้เข้าสอบบางคนโดยไม่ซ้ำกัน กำหนดให้ผู้ตรวจ 1 คน ตรวจให้คะแนนผู้เข้าสอบ 10 คน

6. สัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) หมายถึง ค่าเฉลี่ยของค่าสหสัมพันธ์ระหว่างชุดเงื่อนไขการวัด ที่สุ่มมาจากเอกภพที่ต้องการสรุปอ้างอิง ค่าสัมประสิทธิ์การสรุปอ้างอิงคำนวณได้จากสัดส่วนความแปรปรวนของคะแนนเอกภพต่อค่าความคาดหวังความแปรปรวนของคะแนนสังเกต

7. ความเชื่อมั่นของผู้ตรวจ (Reliability of Rater) หมายถึง ค่าความสอดคล้องของผลการตรวจให้คะแนนของผู้ตรวจจำนวน 5 คน โดยพิจารณาจากค่าเฉลี่ยของผู้ตรวจ ค่าความเชื่อมั่นรวมที่ปรับขยายโดยใช้สูตรของ สเปียร์แมน – บราวน์ และค่าความคลาดเคลื่อนมาตรฐานของการวัด

8. ผู้ตรวจ (Rater) หมายถึง อาจารย์ที่สอนวิชาวิทยาศาสตร์ในระดับมัธยมศึกษาตอนปลายโรงเรียนมัธยมศึกษาสังกัดกรมสามัญศึกษาจังหวัดนครราชสีมา จำนวน 5 ท่าน มีประสบการณ์ในการสอน 5 ปี ขึ้นไป

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องในการศึกษาครั้งนี้ได้แบ่งออกเป็นหัวข้อตามลำดับ ดังนี้

1. ทักษะกระบวนการทางวิทยาศาสตร์
 - 1.1 ความหมายทักษะกระบวนการทางวิทยาศาสตร์
 - 1.2 ประเภททักษะกระบวนการทางวิทยาศาสตร์
 - 1.3 ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ
 - 1.4 การประเมินผลทักษะกระบวนการทางวิทยาศาสตร์
2. แบบทดสอบความเรียง
 - 2.1 ความหมายของแบบทดสอบความเรียง
 - 2.2 วิธีการสร้างแบบทดสอบความเรียง
 - 2.3 อิทธิพลที่มีต่อการวัดผลโดยใช้แบบทดสอบความเรียง
3. กฎเกณฑ์การให้คะแนน (Scoring Rubric)
 - 3.1 ความหมายของกฎเกณฑ์การให้คะแนน
 - 3.2 ประเภทของกฎเกณฑ์การให้คะแนน
 - 3.3 การสร้างกฎเกณฑ์การให้คะแนน
 - 3.4 การตรวจให้คะแนน
4. ทฤษฎีค่าความเชื่อมั่น
 - 4.1 การประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม
 - 4.2 การประมาณค่าความเชื่อมั่นที่เกี่ยวกับการตรวจให้คะแนน
 - 4.3 การประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง
5. งานวิจัยที่เกี่ยวข้อง

1. ทักษะกระบวนการทางวิทยาศาสตร์ (Science Process Skills)

1.1 ความหมายทักษะกระบวนการทางวิทยาศาสตร์

คำว่า "ทักษะกระบวนการทางวิทยาศาสตร์ (Science Process Skills)" ได้มีนักการศึกษาให้ความหมายไว้หลายทัศนะดังนี้

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2524 : 1) ให้ความหมายทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า "เป็นพฤติกรรมที่เกิดจากการปฏิบัติและการฝึกฝนความคิดอย่างมีระบบ ซึ่งก่อให้เกิดความงอกงามทางสติปัญญา"

นิคม ทาแดง และสุจินต์ วิศวธีรานนท์ (2529 : 48) ให้ความหมายของทักษะกระบวนการทางวิทยาศาสตร์เป็นองค์ประกอบที่สำคัญประการหนึ่งของการแสวงหาความรู้ทางวิทยาศาสตร์ เพราะการทำงานตามขั้นตอนของวิธีการทางวิทยาศาสตร์แต่ละขั้นตอนนั้น จะประสบผลสำเร็จหรือล้มเหลว ขึ้นอยู่กับความสามารถและทักษะกระบวนการทางวิทยาศาสตร์ของแต่ละคน ทักษะกระบวนการทางวิทยาศาสตร์สามารถฝึกให้ดีขึ้นได้

วรรณทิพา รอดแรงคำ และพิมพ์พันธ์ เตชะคุปต์ (2532 : V) ให้ความหมายของทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า “เป็นทักษะทางสติปัญญาที่นักวิทยาศาสตร์และผู้ที่เกี่ยวข้องนำวิธีการทางวิทยาศาสตร์มาแก้ปัญหา ใช้ในการศึกษาค้นคว้า สืบเสาะหาความรู้และแก้ปัญหาต่าง ๆ”

คลอฟเฟอร์ (Klopfer. 1971 : 568-573) ให้ความหมายของทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า “ทักษะกระบวนการทางวิทยาศาสตร์เป็นกระบวนการที่ใช้ในการแสวงหาความรู้ทางวิทยาศาสตร์”

ปีเตอร์สัน (Peterson. 1978 : 153) ได้ให้ความหมายของทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า “เป็นปฏิบัติการสืบเสาะแสวงหาความรู้ทางวิทยาศาสตร์ที่ประกอบด้วย การสังเกต การตั้งคำถาม การทดลอง การเปรียบเทียบ การสรุปพาดพิง การสรุปหลักเกณฑ์ การสื่อความหมายและการนำไปใช้”

แอนเดอร์สัน (Anderson. 1978 : 15) ได้ให้ความหมายของทักษะกระบวนการทางวิทยาศาสตร์ไว้ว่า “เป็นวิธีการที่นักวิทยาศาสตร์ใช้ในการเสาะแสวงหาความรู้ ความหมายที่สำคัญของกระบวนการ คือ เป็นวิถีทางของกระบวนการในการเสาะแสวงหาความรู้ ทำให้เกิดพัฒนาการทางด้านสติปัญญา”

จากความหมายและแนวคิดเกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์ที่ได้กล่าวไว้สรุปได้ว่า ทักษะกระบวนการทางวิทยาศาสตร์ หมายถึง ความสามารถในการใช้กระบวนการทางวิทยาศาสตร์ต่าง ๆ ได้แก่ การสังเกต การจำแนกประเภท การวัด การคำนวณ การหาความสัมพันธ์ระหว่างสเปสและเวลา การจัดกระทำและสื่อความหมายข้อมูล การลงความเห็นจากข้อมูล การพยากรณ์ การตั้งสมมุติฐาน การกำหนดและควบคุมตัวแปร การกำหนดนิยามเชิงปฏิบัติการ การตีความหมายข้อมูลและลงข้อสรุป และการทดลองได้อย่างคล่องแคล่วถูกต้องและแม่นยำ เพื่อการเสาะแสวงหาความรู้ หรือแก้ปัญหาอันเกิดจากการปฏิบัติและฝึกฝนความคิดอย่างเป็นระบบ

1.2 ประเภทของทักษะกระบวนการทางวิทยาศาสตร์

หน่วยงานทางการศึกษาและนักการศึกษาได้จัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์ไว้ดังนี้

ในปี พ.ศ. 2519 สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2519 : 8-9) ได้จัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 9 ทักษะ คือ

1. ทักษะการสังเกต
2. ทักษะในการเลือกใช้เครื่องมือ
3. ทักษะในการบันทึกข้อมูลและสื่อความหมาย
4. ทักษะในการจัดกระทำข้อมูล
5. ทักษะในการแปลความหมายข้อมูลและลงข้อสรุป
6. ทักษะในการสร้างสมมุติฐาน
7. ทักษะในการออกแบบการทดลอง
8. ทักษะในการคิดคำนวณ
9. ทักษะในการหาความสัมพันธ์ระหว่างมิติ

ในปี พ.ศ. 2524 สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2524 : 1-16) ได้จัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 2 ประเภท 13 ทักษะ คือ

1. ทักษะกระบวนการขั้นพื้นฐาน จำนวน 8 ทักษะ ประกอบด้วย
 - 1.1 ทักษะการสังเกต
 - 1.2 ทักษะการจำแนกประเภท
 - 1.3 ทักษะการวัด
 - 1.4 ทักษะการคำนวณ
 - 1.5 ทักษะการหาความสัมพันธ์ระหว่างสเปกกับสเปสและสเปกกับเวลา
 - 1.6 ทักษะการจัดกระทำข้อมูลและสื่อความหมาย
 - 1.7 ทักษะการลงความเห็นจากข้อมูล
 - 1.8 ทักษะการพยากรณ์
2. ทักษะขั้นผสมหรือบูรณาการ จำนวน 5 ทักษะ ประกอบด้วย
 - 2.1 ทักษะการตั้งสมมุติฐาน
 - 2.2 ทักษะการกำหนดและควบคุมตัวแปร
 - 2.3 ทักษะการกำหนดนิยามเชิงปฏิบัติการ
 - 2.4 ทักษะการตีความหมายข้อมูลและลงข้อสรุป
 - 2.5 ทักษะการทดลอง

ทบวงมหาวิทยาลัย (2525 : 59) ได้จัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 13 ทักษะ คือ

1. ทักษะการสังเกต
2. ทักษะการจำแนกประเภท
3. ทักษะการวัด
4. ทักษะการคำนวณ
5. ทักษะการหาความสัมพันธ์ระหว่างสเปกกับสเปสและสเปกกับเวลา
6. ทักษะการจัดกระทำข้อมูลและสื่อความหมาย
7. ทักษะการลงความเห็นจากข้อมูล
8. ทักษะการพยากรณ์

9. ทักษะการตั้งสมมุติฐาน
10. ทักษะการกำหนดและควบคุมตัวแปร
11. ทักษะการกำหนดนิยามเชิงปฏิบัติการ
12. ทักษะการตีความหมายข้อมูลและลงข้อสรุป
13. ทักษะการทดลอง

ซันด์ และโทรวบริดจ์ (Sund and Trowbridge. 1967 : 93-95) ได้จัดประเภททักษะกระบวนการทางวิทยาศาสตร์ที่ผู้เรียนควรได้รับการพัฒนาออกเป็น 5 กลุ่ม คือ

1. ทักษะในการหาความรู้ (Acquisitive Skills)
 - 1.1 การฟังอย่างตั้งใจ กระตือรือร้นและถามเมื่อสงสัย
 - 1.2 การสังเกตอย่างถี่ถ้วน สนใจและคิดอย่างเป็นระบบ
 - 1.3 การค้นหาแหล่งข้อมูล ใช้แหล่งข้อมูลหลายแห่งมารวมกัน
 - 1.4 การแสวงหาความรู้โดยการสัมภาษณ์ หรือส่งแบบสอบถาม
 - 1.5 การกำหนดปัญหา
 - 1.6 การเก็บรวบรวมข้อมูลโดยการบันทึกเป็นตาราง หรือจำแนกเป็นรายการต่าง ๆ
 - 1.7 การหาคำตอบของปัญหา โดยทำการทดลอง วิเคราะห์ผลการทดลองแล้วสรุปผล
2. ทักษะในการรวบรวมประสบการณ์ (Organizational Skills)
 - 2.1 การรายงานข้อมูลอย่างมีระบบและสมบูรณ์
 - 2.2 การเปรียบเทียบความเหมือนกันของสิ่งที่สังเกตได้
 - 2.3 การเปรียบเทียบความแตกต่างของสิ่งที่สังเกตได้
 - 2.4 การจำแนกข้อมูลเป็นหมวดหมู่
 - 2.5 การเรียงลำดับข้อมูลที่จัดไว้เป็นหมวดหมู่โดยแสดงเป็นลำดับ
 - 2.6 การกำหนดเค้าโครงออกเป็นหัวข้อใหญ่และหัวข้อย่อย
 - 2.7 การแสดงหัวข้อที่สำคัญและความสัมพันธ์ของข้อมูล
 - 2.8 การประเมินผลและปรับปรุงแก้ไข
 - 2.9 การวิเคราะห์และนำผลไปใช้
3. ทักษะในการสร้างสรรค์ (Creative Skills)
 - 3.1 การวางแผนล่วงหน้าโดยเล็งเห็นผลที่จะเป็นไปได้ และรวมถึงการตั้งสมมุติฐาน
 - 3.2 การกำหนดปัญหาใหม่ วิธีการใหม่ เครื่องมือใหม่และระบบใหม่
 - 3.3 การค้นหาเทคนิควิธีการต่าง ๆ
 - 3.4 การสังเคราะห์โดยการนำสิ่งต่าง ๆ ที่มีอยู่มาประกอบกันเป็นสิ่งใหม่
4. ทักษะในการใช้เครื่องมือ (Manipulative Skills)
 - 4.1 การรู้จักส่วนต่าง ๆ ของเครื่องมือ วิธีการทำงาน การปรับเครื่องมือให้เหมาะสมและข้อจำกัดของงาน
 - 4.2 การดูแลรักษาเครื่องมือให้อยู่ในสภาพดี โดยการเก็บและใช้อย่างเหมาะสม
 - 4.3 การสาธิตแสดงส่วนต่าง ๆ หน้าที่และการทำงานของเครื่องมือ
 - 4.4 การนำเครื่องมือมาใช้ในการทดลอง การวางแผนใช้เครื่องมือ การรวบรวมข้อมูล การบันทึกข้อมูล การวิเคราะห์ข้อมูลและสรุปผลที่ได้จากการใช้เครื่องมือนั้น ๆ

- 4.5 การซ่อมแซมเครื่องมือ
- 4.6 การสร้างเครื่องมืออย่างง่าย ๆ เพื่อใช้ในการทดลอง
- 4.7 การวัดโดยใช้เครื่องมือต่าง ๆ เช่น เทอร์โมมิเตอร์ ตาชั่ง เป็นต้น

5. ทักษะในการสื่อความหมาย (Communicative Skills)

- 5.1 การตั้งคำถาม รู้จักเลือกใช้คำถามที่ดี เป็นคำถามที่ส่งเสริมให้ค้นคว้าหาคำตอบด้วยตนเอง
- 5.2 การอภิปราย รู้จักใช้ความคิดของตนเองและรับฟังความคิดเห็นของผู้อื่น อภิปรายเรื่องที่ตรงประเด็น
- 5.3 การอภิปรายเน้นสาระสำคัญให้ชัดเจน
- 5.4 การรายงานด้วยปากเปล่าต่อชั้นเรียนหรือครู โดยเน้นเนื้อหาสาระที่สำคัญทางวิทยาศาสตร์
- 5.5 การเขียนรายงานการทดลอง หรือการสาธิต เพื่อชี้แจงวิธีการทดลอง การรวบรวมข้อมูล การวิเคราะห์ข้อมูลและสรุปผลที่ได้
- 5.6 การพิจารณาสร้างสรรค์เพื่อประเมินค่าจากผลที่ได้
- 5.7 การเขียนกราฟแสดงผลการทดลองและแปลความหมายจากกราฟ
- 5.8 สามารถถ่ายทอดความรู้ที่ได้แก่เพื่อนร่วมชั้นเรียน

สมาคมอเมริกันเพื่อความก้าวหน้าทางวิทยาศาสตร์ (American Association for the Advancement of Science: AAAS. 1970: 11-176) ได้วิเคราะห์การทำงานของนักวิทยาศาสตร์ และจัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 2 ประเภท 13 ทักษะ คือ

1. ทักษะกระบวนการขั้นพื้นฐาน จำนวน 8 ทักษะ ประกอบด้วย
 - 1.1 ทักษะการสังเกต (Observing)
 - 1.2 ทักษะการใช้ความสัมพันธ์ระหว่างสเปสและเวลา (Using Space and Relationships)
 - 1.3 ทักษะการจัดจำพวก (Classifying)
 - 1.4 ทักษะการใช้ตัวเลข (Using Numbers)
 - 1.5 ทักษะการวัด (Measuring)
 - 1.6 ทักษะการสื่อความหมาย (Communicating)
 - 1.7 ทักษะการทำนาย (Predicting)
 - 1.8 ทักษะการสรุปอ้างอิง (Inferring)
2. ทักษะกระบวนการขั้นบูรณาการหรือขั้นผสม จำนวน 5 ทักษะ ประกอบด้วย
 - 2.1 ทักษะการกำหนดและควบคุมตัวแปร (Identifying controlling and Manipulating Variables)
 - 2.2 ทักษะการประมวลผลและตีความหมายข้อมูล (Data Processing and Interpreting)
 - 2.3 ทักษะการตั้งสมมุติฐาน (Formulating Hypothesis)
 - 2.4 ทักษะการให้นิยามเชิงปฏิบัติการ (Defining Operationally)
 - 2.5 ทักษะการออกแบบและดำเนินการทดลอง (Designing Investigations and Experimenting)

คลอปเฟอร์ (Klopfer, 1971 : 568-573) ได้จัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์ไว้

4 ทักษะ คือ

1. ทักษะการสังเกตและการวัด (Observing and Measuring)
2. ทักษะการมองเห็นปัญหาและวิธีการหาคำตอบ (Seeing a Problem and Seeking Ways to Solve It)
3. ทักษะการแปลความหมายข้อมูลและการสรุป (Interpreting Data and Formulating Generalizations)
4. ทักษะการสร้างทฤษฎี การตรวจสอบและการปรับปรุงทฤษฎีที่สร้างขึ้น (Building, Testing and Revising a Theoretical Model)

โอเคย์ และฟิล (Okey and Field, 1973 : 1-10) ได้จัดประเภททักษะกระบวนการทางวิทยาศาสตร์ที่จำเป็นสำหรับครูผู้สอนวิทยาศาสตร์ไว้ 10 ทักษะ คือ

1. ทักษะการกำหนดตัวแปร (Identifying Variables)
2. ทักษะการสร้างตารางเสนอข้อมูล (Constructing a Table of Data)
3. ทักษะการเขียนกราฟ (Constructing a Graph)
4. ทักษะการอธิบายความสัมพันธ์ของตัวแปร (Describing Relationships between Variables)
5. ทักษะการรวบรวมและจัดกระทำข้อมูล (Acquiring and Processing Your Own Data)
6. ทักษะการวิเคราะห์กระบวนการทดลอง (Analyzing Investigations)
7. ทักษะการตั้งสมมุติฐาน (Constructing Hypothesis)
8. ทักษะการให้นิยามเชิงปฏิบัติการ (Defining Variables Operationally)
9. ทักษะการออกแบบการทดลอง (Designing Investigations)
10. ทักษะการดำเนินการทดลอง (Experimenting)

การ์แลนด์ และคณะ (Garland and Others, 1973 : 127) ได้จัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 9 ทักษะ คือ

1. ทักษะการสังเกต (Observing)
2. ทักษะการจัดกระทำข้อมูล (Data Treatment)
3. ทักษะการพยากรณ์และตั้งสมมุติฐาน (Predicting and Formulating Hypothesis)
4. ทักษะการจำแนกประเภท (Classifying)
5. ทักษะการบ่งชี้ (Identifying)
6. ทักษะการวัด (Measuring)
7. ทักษะการพัฒนาเทคนิควิธีการปฏิบัติในห้องทดลอง (Development of Acceptable Laboratory Technique)
8. ทักษะการวิเคราะห์และสังเคราะห์ (Analysis and Synthesis)
9. ทักษะการสื่อความหมาย (Communicating)

มาร์แชลล์ และ เฮลส์ (Marshall and Hales. 1971 : 105) ได้จัดประเภททักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 5 ประเภท 17 ทักษะ คือ

1. ความคิดริเริ่ม (Initiation)
 - 1.1 การกำหนดขอบเขตของปัญหา
 - 1.2 การหาข้อมูลขั้นต้น
 - 1.3 การทำนาย
 - 1.4 การตั้งสมมุติฐาน
 - 1.5 การเก็บรวบรวมข้อมูลเพื่อการทดลอง
2. การเก็บรวบรวมข้อมูล (Collection of Data)
 - 2.1 การปฏิบัติ
 - 2.2 การสังเกต
3. การจัดกระทำกับข้อมูล (Processing of Data)
 - 3.1 การจัดข้อมูล
 - 3.2 การแสดงข้อมูลโดยกราฟ
 - 3.3 การจัดกระทำข้อมูลในทางคณิตศาสตร์
4. การสร้างมโนทัศน์จากข้อมูล (Conceptualization of Data)
 - 4.1 การตีความหมายข้อมูล
 - 4.2 การสร้างนิยามเชิงปฏิบัติการ
 - 4.3 การแสดงข้อมูลในรูปของความสัมพันธ์ทางคณิตศาสตร์
 - 4.4 การเชื่อมโยงข้อค้นพบใหม่กับทฤษฎีที่มีอยู่
5. ปลายเปิด (Open end)
 - 5.1 ค้นหาหลักฐานต่อไป
 - 5.2 การระบุปัญหาใหม่เพื่อสืบสอบความรู้
 - 5.3 การนำความรู้ที่ค้นพบไปประยุกต์ใช้

โชวาลเตอร์ (Showalter. 1974 : 1-8) ได้จัดประเภททักษะกระบวนการทางวิทยาศาสตร์ต่าง ๆ ที่สำคัญ โดยอาศัยแนวคิดของคณะกรรมการการศึกษาวิทยาศาสตร์ ของสมาคมอเมริกันเพื่อความก้าวหน้าทางวิทยาศาสตร์ (AAAS) ไว้ 15 ทักษะ คือ

1. ทักษะการจำแนก (Classifying)
2. ทักษะการสื่อความหมาย (Communicating)
3. ทักษะการควบคุมตัวแปร (Controlling Variables)
4. ทักษะการให้คำนิยาม (Defining Operationally)
5. ทักษะการออกแบบการทดลอง (Designing Experimenting)
6. ทักษะการสร้างแบบจำลอง (Formulating Models)
7. ทักษะการตั้งสมมุติฐาน (Formulating Hypothesis)
8. ทักษะการลงความคิดเห็น (Inferring)
9. ทักษะการตีความหมายข้อมูล (Interpreting Data)
10. ทักษะการวัด (Measuring)

11. ทักษะการสังเกต (Observing)
12. ทักษะการทำนาย (Predicting)
13. ทักษะการตั้งคำถาม (Questioning)
14. ทักษะการใช้ตัวเลข (Using Numbers)
15. ทักษะการใช้ความสัมพันธ์เกี่ยวกับช่วงเวลา (Using Spacetime Relationships)

ลันสตรอม และโลเวอรี (Lunstrom and Lowery. 1978 : 209-210) ได้จัดประเภทของทักษะกระบวนการทางวิทยาศาสตร์เป็น 10 ทักษะ คือ

1. ทักษะการสังเกตและการบรรยาย (Observation and Description)
2. ทักษะการเปรียบเทียบ (Comparison)
3. ทักษะการจำแนก (Classification)
4. ทักษะการใช้เทคนิคทางคณิตศาสตร์ (Mathematical Techniques)
5. ทักษะการปฏิบัติการทดลอง (Experimentation)
6. ทักษะการสรุปอ้างอิง (Inferences)
7. ทักษะการแสดงความหมาย (Implications)
8. ทักษะการทดลองเพื่อทดสอบสมมุติฐาน (Verification in Independent Study)
9. ทักษะการศึกษาอย่างลึกซึ้ง (Depth Study)
10. ทักษะการใช้เครื่องมือและความสามารถเฉพาะ (Instrumentation and Other Specialization)

ดูเรน (Duren. 1978 : 20) กล่าวว่า สำนักประเมินผลการศึกษาเพื่อความก้าวหน้าแห่งชาติ (The National Assessment of Education Progress : NAEP, 1978) ได้แบ่งทักษะกระบวนการทางวิทยาศาสตร์ที่จำเป็นสำหรับการทำงานทางวิทยาศาสตร์ ออกเป็น 10 ทักษะ คือ

1. ทักษะการให้นิยามทางวิทยาศาสตร์ (Define a Scientific Problem)
2. ทักษะการเสนอแนะหรือรู้สมมุติฐานทางวิทยาศาสตร์ (Suggest or Recognize a Scientific Hypothesis)
3. ทักษะการเสนอหรือเลือกวิธีการที่เหมาะสมในด้านเหตุผลและการปฏิบัติ (Propose or Select Validating Procedures)
4. ทักษะการเก็บข้อมูล (Obtain Requisite Data)
5. ทักษะการตีความหมายข้อมูล (Interpret Data)
6. ทักษะการตรวจสอบความถูกต้องอย่างมีเหตุผลของสมมุติฐานกับกฎ ข้อเท็จจริง การสังเกต หรือการทดลอง (Check the Logical Consistency of Hypothesis with Relevant Laws, Fact, Observations or Experiments)
7. ทักษะการให้เหตุผลทางด้านปริมาณและสัญลักษณ์ (Reason Quantitative and Symbolically)
8. ทักษะการจำแนกความแตกต่างระหว่างข้อเท็จจริง สมมุติฐานและความคิดเห็น สิ่งที่เกี่ยวข้องกับสิ่งที่ไม่เกี่ยวข้องและรูปแบบจากการสังเกต (Distinguish Between Fact, Hypothesis and Opinions the Relevant from the Irrelevant and the Model from the Observations)
9. ทักษะการวิเคราะห์และวิจารณ์เอกสารทางวิทยาศาสตร์ (Read Scientific Materials Critically)

10. ทักษะการใช้กฎและหลักการทางวิทยาศาสตร์ ทั้งในสถานที่คุ้นเคยและไม่คุ้นเคย (Employ Scientific Laws and Principles in Familiar and Unfamiliar Situations)

เจคอบสัน และ บาร์รี (Jacobson and Barry, 1980 : 450-451) ได้จัดประเภททักษะกระบวนการทางวิทยาศาสตร์ที่ควรฝึกให้กับนักเรียน เพื่อให้นักเรียนได้มีโอกาสพัฒนาความสามารถด้านนี้ และสามารถนำเอาทักษะด้านนี้ไปใช้ในชีวิตประจำวันได้ โดยแบ่งทักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 13 ทักษะ คือ

1. ทักษะการกำหนดปัญหา (Defining Problems)
2. ทักษะการวางแผนทางในการค้นคว้า (Designing Investigations)
3. ทักษะการทดลองที่มีการควบคุม (Controlled Experiments)
4. ทักษะการสังเกต (Observing)
5. ทักษะการวัด (Measuring)
6. ทักษะการจำแนก (Classifying)
7. ทักษะการสรุปลงความคิดเห็น (Inferring)
8. ทักษะการตั้งสมมุติฐาน (Formulating Hypothesis)
9. ทักษะการรวบรวมและจัดกระทำข้อมูล (Collecting and Handling Data)
10. ทักษะการแปลความหมายข้อมูล (Interpreting Data)
11. ทักษะการค้นคว้าเอกสารที่เกี่ยวข้อง (Literature Search)
12. ทักษะการสร้างความสัมพันธ์กับทฤษฎี (Relating to Theory)
13. ทักษะการลงข้อสรุป (Drawing Conclusions)

จากแนวคิดการจัดประเภทกระบวนการทางวิทยาศาสตร์ที่กล่าวมา ผู้วิจัยขอจัดประเภททักษะกระบวนการทางวิทยาศาสตร์ออกเป็น 2 ประเภท 13 ทักษะ คือ

1. ทักษะกระบวนการขั้นพื้นฐาน (Basic Science Process Skills) เป็นทักษะกระบวนการที่ผู้เรียนควรฝึกฝนให้เกิดความชำนาญเป็นพื้นฐาน ก่อนไปฝึกทักษะกระบวนการขั้นผสม หรือขั้นบูรณาการ ทักษะกระบวนการขั้นพื้นฐานประกอบด้วย 8 ทักษะ คือ

- 1.1 ทักษะการสังเกต (Observing)
- 1.2 ทักษะการจำแนกประเภท (Classifying)
- 1.3 ทักษะการวัด (Measuring)
- 1.4 ทักษะการคำนวณ (Using Numbers)
- 1.5 ทักษะการหาความสัมพันธ์ระหว่างสเปซกับสเปสและสเปซกับเวลา (Space/Space and Space/Time Relationships)
- 1.6 ทักษะการจัดกระทำข้อมูลและสื่อความหมาย (Organizing Data and Communicating)
- 1.7 ทักษะการลงความเห็นจากข้อมูล (Inferring)
- 1.8 ทักษะการพยากรณ์ (predicting)

2. ทักษะกระบวนการขั้นผสมหรือขั้นบูรณาการ (Integrated Science Process Skills) เป็นทักษะกระบวนการที่ต้องอาศัยการบูรณาการจากทักษะกระบวนการขั้นพื้นฐาน ซึ่งผู้เรียนควรฝึกฝนให้เกิดความชำนาญมาก่อน จึงจะทำให้ทักษะกระบวนการขั้นผสมหรือขั้นบูรณาการมีประสิทธิภาพ ประกอบด้วย 5 ทักษะ คือ

- 2.1 ทักษะการตั้งสมมุติฐาน (Formulating Hypothesis)
- 2.2 ทักษะการกำหนดนิยามเชิงปฏิบัติการ (Defining Operationally)
- 2.3 ทักษะการกำหนดและควบคุมตัวแปร (Identifying and Controlling Variables)
- 2.4 ทักษะการทดลอง (Experimenting)
- 2.5 ทักษะการตีความหมายข้อมูลและลงข้อสรุป (Interpreting Data and Making

Conclusion)

ชุดิญา สุริยมณฑล (2535 : 68, 74-75) ได้สำรวจประเภททักษะกระบวนการทางวิทยาศาสตร์ ในหนังสือเรียนวิทยาศาสตร์ระดับมัธยมศึกษาตอนต้นจำนวน 6 เล่ม โดยใช้ทักษะกระบวนการทางวิทยาศาสตร์ของสมาคมอเมริกันเพื่อความก้าวหน้าในทางวิทยาศาสตร์ (AAAS) 13 ทักษะเป็นเกณฑ์ ผลการสำรวจพบว่า ในหนังสือเรียนวิทยาศาสตร์ระดับมัธยมศึกษาตอนต้นทุกเล่มมีทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน ร้อยละ 78.55 และทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ร้อยละ 21.45

จากการสำรวจดังกล่าว พบว่า ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการมีสัดส่วนที่ประกอบอยู่ในหลักสูตรวิทยาศาสตร์ชั้นมัธยมศึกษาตอนปลาย ทักษะกระบวนการทางวิทยาศาสตร์ขั้นนี้ ถือเป็นทักษะกระบวนการขั้นสูง เป็นทักษะที่กำหนดไว้ในระดับมัธยมศึกษาตอนปลาย เพื่อเป็นแนวทางในการศึกษาต่อและนำไปใช้ในชีวิตประจำวัน ผู้วิจัยจึงเลือกที่จะศึกษา แนวทางในการประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ โดยการพัฒนากฎเกณฑ์การให้คะแนน (Scoring Rubric) ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ที่บ่งชี้คุณภาพ โดยอาศัยความสอดคล้องของการตรวจให้คะแนนของผู้ตรวจ ค่าความเชื่อมั่นของผู้ตรวจและค่าความคลาดเคลื่อนมาตรฐานของการวัด พร้อมทั้งศึกษาการลดภาระของการตรวจให้คะแนนโดยใช้ค่าสัมประสิทธิ์การสุปอ้างอิง ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการทั้ง 5 ด้าน มีรายละเอียดดังนี้

1.3 ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

1.3.1 ทักษะการตั้งสมมุติฐาน

จากการศึกษาค้นคว้า ได้มีนักการศึกษาได้ให้ความหมายของทักษะการตั้งสมมุติฐานไว้หลายท่าน ดังนี้

ปรีชา วงศ์ชูศิริ และคณะ (2525 : 349) ให้ความหมายว่า ทักษะการตั้งสมมุติฐานหมายถึงความสามารถในการให้ข้อสรุปหรือคำอธิบาย เป็นคำตอบล่วงหน้าก่อนที่จะดำเนินการทดลอง เพื่อตรวจสอบความถูกต้องเป็นจริงของเรื่องนั้น ๆ ต่อไป ตัวสมมุติฐานเป็นข้อความที่แสดงความสัมพันธ์ที่เชื่อได้ว่า จะเกิดขึ้นระหว่างตัวแปรอิสระกับตัวแปรตาม ข้อความของสมมุติฐานถูกสร้างขึ้นโดยอาศัยการสังเกต ความรู้และประสบการณ์เดิม กฎ หลักหรือทฤษฎีที่เกี่ยวข้อง ข้อความใดจะจัดว่าเป็นสมมุติฐานก็ต่อเมื่อข้อความ

นั้นอ้างถึงข้อเท็จจริง ที่ยังไม่เคยมีประสบการณ์มาก่อน สามารถตรวจสอบโดยการทดลองได้และแก้ไขเมื่อมีความรู้ใหม่

ภพ เลหาไฟบูลย์ (2537 : 25) กล่าวว่า ทักษะการตั้งสมมุติฐาน หมายถึง ความสามารถในการให้คำอธิบายซึ่งเป็นคำตอบล่วงหน้าก่อนที่จะดำเนินการทดลอง เพื่อตรวจสอบความถูกต้องเป็นจริงในเรื่องนั้น ๆ ต่อไป สมมุติฐานเป็นข้อความที่แสดงการคาดคะเน ซึ่งอาจเป็นคำอธิบายของสิ่งที่ไม่สามารถตรวจสอบโดยการสังเกตได้ หรืออาจเป็นข้อความที่แสดงความสัมพันธ์จากการสังเกต ความรู้ ประสบการณ์เดิมเป็นพื้นฐาน การคาดคะเนคำตอบล่วงหน้ายังไม่ทราบหรือยังไม่เป็นหลักการ กฎหรือทฤษฎีมาก่อน ข้อความของสมมุติฐานต้องสามารถทำการตรวจสอบโดยการทดลองและแก้ไขเมื่อมีความรู้ใหม่ได้

สมาคมอเมริกันเพื่อความก้าวหน้าทางวิทยาศาสตร์ (AAAS. 1970) ให้ความหมายของทักษะการตั้งสมมุติฐานไว้ว่า "เป็นความสามารถในการคาดการณ์ว่า ตัวแปรต่าง ๆ จะมีความสัมพันธ์กันอย่างไร เป็นการลงข้อสรุปของคำอธิบาย โดยอาศัยการสังเกตหรือการสรุปอ้างอิงเป็นพื้นฐาน"

กูต (Good. 1973 : 289) ให้ความหมายของทักษะการตั้งสมมุติฐานไว้ว่า "เป็นความสามารถในการคาดคะเนคำตอบเกี่ยวกับการแก้ปัญหา หรือกฎทั่วไปที่จะอธิบายองค์ประกอบต่าง ๆ"

จากความหมายหรือแนวคิดดังกล่าว พอสรุปได้ว่า ทักษะการตั้งสมมุติฐาน หมายถึง ความสามารถในการให้ข้อสรุปหรือการอธิบายถึงความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ว่ามีความสัมพันธ์กันอย่างไร หรือเป็นการคิดหาคำตอบล่วงหน้าก่อนที่จะดำเนินการทดลอง เพื่อตรวจสอบความถูกต้องในเรื่องนั้น ๆ สมมุติฐานสร้างขึ้นโดยอาศัยการสังเกต ความรู้ และประสบการณ์เดิมเป็นพื้นฐาน สมมุติฐานต้องเป็นสิ่งที่ยังไม่ทราบหรือมีประสบการณ์มาก่อนหรือยังไม่เป็นหลักการ กฎหรือทฤษฎีมาก่อน เราสามารถตรวจสอบได้โดยการทดลองและแก้ไขได้เมื่อมีความรู้ใหม่เกิดขึ้น

สมมุติฐานเป็นเครื่องกำหนดแนวทางในการออกแบบการทดลองเพื่อตรวจสอบว่าสมมุติฐานที่ตั้งขึ้นจะยอมรับหรือไม่ยอมรับ สมมุติฐานที่ตั้งขึ้นอาจจะถูกหรือผิดก็ได้ ซึ่งจะทราบภายหลังการทดลองหาคำตอบ และ ในสถานการณ์การทดลองหนึ่งอาจมี 1 สมมุติฐาน หรือหลายสมมุติฐานก็ได้

สมมุติฐานแตกต่างจากการพยากรณ์ตรงที่ว่า การพยากรณ์เป็นการทำนายหรือคาดคะเนคำตอบล่วงหน้า โดยอาศัยข้อมูลที่ได้จากการสังเกต การวัด รวมไปถึงความสัมพันธ์ระหว่างตัวแปรที่ได้ศึกษามาแล้ว หรือปรากฏการณ์ที่เกิดขึ้นซ้ำ ๆ กันผลการทำนายหรือการคาดคะเนคำตอบจะถูกต้องหรือแม่นยำ เป็นผลมาจากการสังเกตอย่างรอบคอบและการวัดที่ ถูกต้อง และจะแตกต่างกับการลงความคิดเห็นจากข้อมูล โดยที่การลงความคิดเห็นจากข้อมูลจะเป็นการอธิบายสิ่งที่ได้จากการสังเกต การวัดหรือการทดลอง หรือเป็นการเพิ่มความคิดเห็นให้กับข้อมูล ส่วนสมมุติฐานจะเป็นการสร้างความสัมพันธ์ระหว่างตัวแปรต่าง ๆ หรือเป็นการทำนายหรือคาดคะเนคำตอบล่วงหน้าก่อนที่จะดำเนินการทดลองหรือทดสอบสมมุติฐานที่ตั้งขึ้น

พฤติกรรมของการตั้งสมมุติฐาน

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2534 : 55) กำหนดพฤติกรรมหรือความสามารถที่แสดงว่าเกิดทักษะการตั้งสมมุติฐาน คือ หาคำตอบล่วงหน้าก่อนการทดลองโดยอาศัยความสังเกต ความรู้และประสบการณ์เดิม

คำถามที่นำไปสู่ทักษะการตั้งสมมุติฐาน

1. ถ้าเปลี่ยนน้ำหนักถ่วงที่มีของหน้มน้อยนักกายกรรม ผลการทดลองที่ได้จะเปลี่ยนแปลงหรือไม่
 2. ถ้านำปลายแหลมของตุ้กดตามทีประกอบแล้ว (แต่ไม่มีขา) วางไว้บนแท่น นักเรียนคิดว่าตุ้กดขาตัวนี้จะยืนได้หรือไม่ เพราะเหตุใด
- จากพฤติกรรมหรือความสามารถดังกล่าว ผู้วิจัยของกำหนดความสามารถของผู้ที่แสดงว่าเกิดทักษะการตั้งสมมุติฐาน คือ
1. หาคำตอบล่วงหน้าก่อนการทดลองโดยอาศัยการสังเกต ความรู้และประสบการณ์เดิมได้
 2. สร้างหรือแสดงให้เห็นวิธีที่สนับสนุนสมมุติฐานและไม่สนับสนุนสมมุติฐาน ออกจากกันได้

คำถามที่นำไปสู่ทักษะการตั้งสมมุติฐาน เช่น

1. การให้ปุ๋ยแก่ฝรั่งพันธุ์กลมสีทางรากและทางลำต้น จะให้ผลต่างกันอย่างไร
2. ไก่ที่กินหนอนกับไก่ที่กินข้าวจะเจริญเติบโตต่างกันอย่างไร
3. จากการทดลองละลายสาร A ในน้ำ 50 cm³ ที่อุณหภูมิต่างกัน ได้ผลการทดลอง ดังนี้

ตาราง 1 ผลการทดลองละลายสาร A ในน้ำ 50 cm³ ที่อุณหภูมิต่างกัน

อุณหภูมิของน้ำ (องศาเซลเซียส)	ปริมาณสาร A ที่ละลายในน้ำ (กรัม)
60	23
70	31
80	39
90	47

ถ้าเพิ่มปริมาณของน้ำ การละลายของสาร A เป็นอย่างไร

1.3.2. ทักษะการกำหนดนิยามเชิงปฏิบัติการ

จากการศึกษาค้นคว้าได้มีนักการศึกษาได้ให้ความหมายของทักษะการกำหนดนิยามเชิงปฏิบัติการไว้หลายท่าน ดังนี้

ทบวงมหาวิทยาลัย (2525 : 78) กล่าวว่า คำนิยามเชิงปฏิบัติการ เป็นความหมายของคำศัพท์เฉพาะ เป็นภาษาง่าย ๆ ชัดเจน ไม่กำกวม ระบุสิ่งที่จะสังเกต หรือระบุการกระทำซึ่งอาจได้จากการวัด ทดสอบหรือการทดลอง

พิศาล สร้อยธูห์ (2529 : 228) กล่าวว่า การกำหนดข้อความเพื่อสื่อความหมายทางวิทยาศาสตร์ ถ้ากำหนดในเชิงที่ผู้ที่เราต้องการจะติดต่อด้วยสามารถปฏิบัติตามได้ ตรวจสอบได้ ก็จะช่วยสื่อความหมายให้เป็นที่เข้าใจตรงกัน และเป็นประโยชน์ในการที่จะทำการทดลองหรือตรวจสอบได้ด้วย

สมาคมอเมริกันเพื่อความก้าวหน้าทางวิทยาศาสตร์ (AAAS. 1970) ให้ความหมายของทักษะการกำหนดนิยามเชิงปฏิบัติการว่า “เป็นการกำหนดความหมายและขอบเขตของคำต่าง ๆ ที่อยู่ในสมมุติฐานที่ต้องการทดลองให้เข้าใจตรงกัน และสามารถสังเกตได้หรือวัดได้”

กูด (Good. 1973 : 168, 398) ให้ความหมายของทักษะการกำหนดนิยามเชิงปฏิบัติการ ไว้ว่า “เป็นความสามารถในการให้ความหมายหรือความสำคัญของคำ แนวคิดหรือปัญหา ในลักษณะที่สังเกต วัด หรือนำมาปฏิบัติได้”

จากแนวคิดและความหมายของทักษะการกำหนดนิยามเชิงปฏิบัติการดังกล่าว สรุปได้ว่า ทักษะการกำหนดนิยามเชิงปฏิบัติการ หมายถึง ความสามารถในการให้ความหมายของคำหรือข้อความเพื่อให้เข้าใจตรงกันและสามารถสังเกต วัดหรือตรวจสอบได้ ในสถานการณ์การทดลองแต่ละการทดลอง การให้นิยามเชิงปฏิบัติการย่อมแตกต่างกันไปตามสถานการณ์ นั้น ๆ และพร้อมที่จะให้ผู้อื่นทำการทดสอบสิ่งของต่าง ๆ ได้โดยการสังเกต วัดหรือลงมือปฏิบัติเพื่อทดสอบว่า เป็นสิ่งเดียวกันที่ให้ความหมายไว้หรือไม่ ในเหตุการณ์เดียวกันอาจให้นิยามได้หลายอย่าง ทั้งนี้ ขึ้นอยู่กับสถานการณ์การทดลองนั้น ๆ

ตัวอย่างของการกำหนดนิยามเชิงปฏิบัติการและนิยามทั่วไป เช่น

1. ก. ออกซิเจน เป็นก๊าซที่ช่วยให้ไฟติด เมื่อนำก้านไม้ขีดไฟที่เหลือแต่ถ่านแดง ๆ แหย่ลงไป ในหลอดทดลองที่มีก๊าซออกซิเจนอยู่ ก้านไม้ขีดไฟจะลุกเป็นเปลวไฟ

ข. ออกซิเจน เป็นธาตุที่มีเลขอะตอม 8 และน้ำหนักอะตอม 16

จากตัวอย่างข้อ 1 ก. จัดเป็นนิยามเชิงปฏิบัติการ เนื่องจากสามารถสังเกตและระบุสถานการณ์ที่สามารถทดสอบได้ ส่วนตัวอย่าง ข้อ 1 ข. จัดเป็นนิยามทั่วไป เนื่องจากไม่สามารถสังเกตและทดสอบได้ถึงแม้จะเข้าใจตรงกัน

2. ก. การสมดุลของวัตถุ หมายถึง เมื่อมีแรงหลายแรงมากระทำต่อวัตถุก่อนเดียวกัน วัตถุจะไม่มีการเลื่อนตำแหน่งหรือเลื่อนตำแหน่งด้วยความเร็วคงที่ และไม่เกิดการหมุน

ข. การสมดุลของวัตถุ เป็นการสมดุลทั้งการเลื่อนตำแหน่ง และการสมดุลต่อการหมุน

จากตัวอย่างข้อ 2 ก. จัดเป็นนิยามเชิงปฏิบัติการ เนื่องจากระบุการกระทำเมื่อมีแรงหลายแรงมากระทำต่อวัตถุ และระบุสิ่งที่สังเกตได้ คือ วัตถุจะไม่มีการเลื่อนตำแหน่งหรือเลื่อนตำแหน่งด้วยความเร็วคงที่และไม่เกิดการหมุน ส่วนตัวอย่างข้อ 2 ข. จัดเป็นนิยามทั่วไป เนื่องจากระบุสิ่งที่สังเกตได้เพียงอย่างเดียว แต่ไม่ระบุการกระทำ

3. ก. คาร์บอนไดออกไซด์ หมายถึง ก๊าซชนิดหนึ่งที่มีสมบัติที่ไม่ช่วยให้ไฟติด ถ้าจุดไฟ แหย่ลงไป ในหลอดทดลองที่มีก๊าซนี้อยู่ ไฟจะดับอย่างรวดเร็ว

ข. คาร์บอนไดออกไซด์ หมายถึง สารประกอบชนิดหนึ่ง ซึ่งแต่ละโมเลกุลของ คาร์บอนไดออกไซด์ประกอบด้วยธาตุคาร์บอน 1 อะตอม และธาตุออกซิเจน 2 อะตอม

จากตัวอย่างข้อ 3 ก. จัดเป็นนิยามเชิงปฏิบัติการ เพราะระบุการกระทำ (จุดไฟแหย่ลงไป ในหลอดทดลอง) และระบุสิ่งที่สังเกตได้คือ ไฟจะดับอย่างรวดเร็ว ส่วนข้อ 3 ข. จัดเป็นนิยามทั่วไป เพราะไม่สามารถสังเกตหรือทดสอบได้

การกำหนดนิยามเชิงปฏิบัติการส่วนใหญ่จะพบในวิชาวิทยาศาสตร์กายภาพ (ฟิสิกส์และเคมี) ซึ่งเราสามารถระบุได้ชัดเจนว่า จะต้องทำอะไรและจะสังเกตอะไร สำหรับในวิชาวิทยาศาสตร์ชีวภาพ นิยามเชิงปฏิบัติการส่วนใหญ่จะอยู่ในลักษณะที่แตกต่างออกไป เช่น

“ก้านใบ หมายถึง ส่วนของใบที่มีลักษณะคล้ายกิ่ง อยู่ระหว่างตัวใบกับลำต้น พบในพืชบางชนิด มีขนาดพอ ๆ กับตัวใบหรือสั้นกว่า” ถ้าได้ศึกษาเกี่ยวกับนิยามเชิงปฏิบัติการของคำว่า “ใบ” กับ “ลำต้น” มาก่อนแล้ว จากนิยามเชิงปฏิบัติการดังกล่าว เราก็สามารถที่จะชี้บ่งได้ทันทีว่าส่วนที่เป็น “ก้านใบ” คือส่วนไหน จะเห็นได้ว่า นิยามเชิงปฏิบัติการที่ให้ระบุสังเกตอะไรได้อย่างชัดเจน แต่ไม่ได้ระบุว่าทำอะไร ซึ่งในที่นี้ไม่มีความจำเป็นแต่อย่างใด

พฤติกรรมทักษะการกำหนดนิยามเชิงปฏิบัติการ

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2534 : 56) กำหนดพฤติกรรมหรือความสามารถที่แสดงว่า เกิดทักษะการกำหนดนิยามเชิงปฏิบัติการ คือ

“กำหนดความหมายและขอบเขตของคำหรือตัวแปรต่าง ๆ ให้สังเกตได้และวัดได้”

คำถามที่นำไปสู่ทักษะการกำหนดนิยามเชิงปฏิบัติการ เช่น

1. จะกำหนดความหมายของสารแม่เหล็กได้อย่างไร เพื่อให้เข้าใจตรงกัน
2. เส้นสมมาตรมีความหมายว่าอย่างไร
3. ลักษณะอย่างไรจึงจะเรียกว่าลายมือแบบกันหอย

จากพฤติกรรมหรือความสามารถดังกล่าว ผู้วิจัยขอกำหนดความสามารถของผู้ที่แสดงว่าเกิดทักษะการกำหนดนิยามเชิงปฏิบัติการ คือ

1. กำหนดความหมายและขอบเขตของคำ หรือตัวแปรต่าง ๆ ให้สังเกตและวัดได้
2. สามารถแยกนิยามเชิงปฏิบัติการออกจากนิยามทั่วไปได้
3. สามารถชี้บ่งตัวแปรหรือคำที่ต้องใช้ในการให้นิยามเชิงปฏิบัติการได้

คำถามที่นำไปสู่ทักษะการกำหนดนิยามเชิงปฏิบัติการ เช่น

1. การเจริญเติบโตในที่นี้หมายความว่าอย่างไร
2. น้ำสะอาดในที่นี้หมายความว่าอย่างไร
3. จะทราบได้อย่างไรว่าของอย่างหนึ่งเป็นฉนวนไฟฟ้า
4. จุดเยือกแข็งในที่นี้หมายความว่าอย่างไร

1.3.3. ทักษะการกำหนดและควบคุมตัวแปร

จากการศึกษาค้นคว้าได้มีนักการศึกษาได้ให้ความหมายของทักษะการกำหนดและ ควบคุมตัวแปรไว้หลายท่าน ดังนี้

สมชัย โกมล และคณะ (2525 : 7) ให้ความหมายของทักษะการกำหนดและควบคุมตัวแปรว่า “เป็นความสามารถในการบ่งชี้ตัวแปรอิสระ ตัวแปรตามและควบคุมตัวแปรอิสระอื่น ๆ ที่เกี่ยวข้อง แต่ยังไม่ต้องการศึกษาในสมมุติฐานหนึ่ง ๆ”

ภพ เลหาไพบูลย์ (2537 : 27) กล่าวว่า ทักษะการกำหนดและควบคุมตัวแปร หมายถึงการบ่งชี้ตัวแปรต้น ตัวแปรตามและตัวแปรที่ต้องควบคุมในสมมุติฐานหนึ่ง การควบคุมตัวแปรนั้นเป็นการควบคุมสิ่งอื่นๆ นอกเหนือจากตัวแปรต้นที่ทำให้ผลการทดลองคลาดเคลื่อน ถ้าหากว่าไม่ควบคุมให้เหมือนกัน

สมาคมอเมริกันเพื่อความก้าวหน้าทางวิทยาศาสตร์ (AAAS. 1970) ให้ความหมายของทักษะการกำหนดและควบคุมตัวแปรว่าเป็นความสามารถในการบ่งชี้ตัวแปรต้น ตัวแปรตามและตัวแปรที่ต้องควบคุมในสมมุติฐานหนึ่ง ๆ

กูต (Good. 1973 : 638) กล่าวว่า ตัวแปร หมายถึง ลักษณะที่เปลี่ยนแปลงจากเงื่อนไขหนึ่งไปยังอีกเงื่อนไขหนึ่ง

ประเภทของตัวแปร

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2534 : 56) ได้แบ่งตัวแปรออกเป็น 3 ประเภท คือ

1. ตัวแปรต้น คือ สิ่งที่เป็นสาเหตุที่ทำให้เกิดผลต่าง ๆ หรือสิ่งที่เราต้องการทดลองดูว่าเป็นสาเหตุที่ก่อให้เกิดผลเช่นนั้นจริงหรือไม่

2. ตัวแปรตาม คือ สิ่งที่เป็นผลเนื่องมาจากตัวแปรต้น เมื่อตัวแปรต้นหรือสิ่งที่เป็นสาเหตุเปลี่ยนไป ตัวแปรตามหรือสิ่งที่เป็นผลจะเปลี่ยนตามไปด้วย

3. ตัวแปรที่ต้องควบคุม คือ สิ่งอื่น ๆ นอกเหนือจากตัวแปรต้นที่มีผลต่อการทดลองด้วย ซึ่งจะต้องควบคุมให้เหมือน ๆ กัน มิเช่นนั้น อาจทำให้ผลการทดลองคลาดเคลื่อน

การกำหนดตัวแปร หมายถึง การชี้บ่งตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุมในสมมุติฐานหนึ่ง ๆ

จากความหมายและประเภทของตัวแปรดังกล่าว ผู้วิจัยขอให้ความหมายของทักษะการกำหนดและควบคุมตัวแปร ประเภทของตัวแปรและตัวอย่างการกำหนดตัวแปร ดังนี้

ทักษะการกำหนดและควบคุมตัวแปร หมายถึง ความสามารถที่จะบ่งชี้ได้ว่าตัวแปรใดเป็นตัวแปรอิสระหรือตัวแปรต้น ตัวแปรใดเป็นตัวแปรตาม และตัวแปรใดเป็นตัวแปรที่ต้องควบคุม การควบคุมตัวแปรเป็นส่วนสำคัญอย่างยิ่งของการทดลอง ทั้งนี้ เพื่อให้ได้ผลสรุปที่ถูกต้องแน่นอนว่า ผลที่เกิดขึ้นนั้นจากตัวแปรที่เราต้องการจะศึกษาหรือไม่ ในสถานการณ์การทดลองหนึ่ง ๆ ผลที่เกิดขึ้นอาจมาจากหลายสาเหตุ จึงมีความจำเป็นต้องควบคุมสิ่งที่เราไม่ต้องการศึกษาให้เหลือเฉพาะตัวแปรที่เราต้องการจะทราบ เพื่อสะดวกในการศึกษากรณีเฉพาะสาเหตุใดสาเหตุหนึ่ง

ตัวแปร หมายถึง คุณลักษณะของสิ่งต่าง ๆ ที่สามารถแปรผันหรือเปลี่ยนแปลงค่าของตัวเองได้เมื่ออยู่ในสถานการณ์ใดสถานการณ์หนึ่ง ซึ่งอาจเปลี่ยนแปลงในเชิงปริมาณหรือเชิงคุณภาพ

การกำหนดตัวแปร หมายถึง การบ่งชี้ตัวแปรต้น (ตัวแปรอิสระ) ตัวแปรตามและ ตัวแปรที่ต้องควบคุม ในสมมุติฐานหนึ่ง ๆ

การควบคุมตัวแปร หมายถึง การควบคุมตัวแปรอื่น ๆ นอกเหนือจากตัวแปรต้น เพื่อให้ผลการทดลองไม่คลาดเคลื่อน หรือคลาดเคลื่อนน้อยที่สุด

การทดลองทางวิทยาศาสตร์ มีความจำเป็นที่จะต้องควบคุมตัวแปร โดยเฉพาะอย่างยิ่งตัวแปรที่เราไม่ต้องการศึกษา แต่มีผลต่อการทดลอง หากไม่มีการควบคุมตัวแปรแล้วจะทำให้ผลการทดลองคลาดเคลื่อนจนไม่อาจสรุปได้ นอกจากนี้ ความรู้ความเข้าใจเกี่ยวกับการควบคุมตัวแปรเป็นพื้นฐานสำคัญในการออกแบบการทดลอง เพื่อทดสอบสมมุติฐานต่าง ๆ อีกด้วย

ในการทดลองโดยทั่วไปนั้น ในบางครั้งการควบคุมตัวแปรอื่น ๆ ที่ไม่ต้องการศึกษาให้คงที่จริง ๆ นั้นอาจทำได้ยาก และตัวแปรนั้นอาจจัดว่าไม่มีผลต่อการทดลอง เช่น ในการทดลองในห้องปฏิบัติการทั่ว ๆ ไป เราถือว่าอุณหภูมิของห้องคงที่ ซึ่งตามความเป็นจริงแล้วอุณหภูมิห้องขณะทดลองแต่ละครั้งอาจจะไม่เท่ากันก็ได้ แต่ความแตกต่างของอุณหภูมิที่มีผลต่อการทดลองมีน้อยมาก เราจึงไม่คำนึงถึงความแตกต่างนี้

พฤติกรรมของทักษะการกำหนดและควบคุมตัวแปร

สถาบันการส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2534 : 56) ได้กำหนด พฤติกรรมหรือความสามารถที่แสดงว่า เกิดทักษะการกำหนดและควบคุมตัวแปร คือ

“ชี้บ่งและกำหนดตัวแปรต้น ตัวแปรตามและตัวแปรที่ต้องควบคุมได้”

คำถามที่นำไปสู่ทักษะการกำหนดและควบคุมตัวแปร เช่น

ในการทดลองเพื่อทดสอบสมมุติฐานว่า “มุมของปลายรางมีผลต่อระยะตกของ ลูกปืน” ใน การทดลองนี้

1. จะต้องจัดอะไรให้ต่างกัน (ถามให้ชี้บ่งตัวแปรต้น)
2. เมื่อจัดสถานการณ์ดังนี้แล้ว สิ่งที่เราจะต้องติดตามดูคืออะไร (ถามให้ชี้บ่งตัวแปรตาม)
3. จะต้องจัดอะไรให้เหมือนกันบ้าง (ถามให้ชี้บ่งตัวแปรที่ต้องควบคุม)

จากพฤติกรรมหรือความสามารถดังกล่าว ผู้วิจัยขอกำหนดความสามารถของผู้ที่แสดงว่าเกิดทักษะการกำหนดและควบคุมตัวแปร ดังนี้

กำหนดตัวแปรต้น ตัวแปรตามและตัวแปรที่ต้องควบคุมได้

คำถามที่นำไปสู่ทักษะการกำหนดและควบคุมตัวแปร เช่น

ในการทดลองเพื่อทดสอบสมมุติฐานว่า “การให้ปุ๋ยแก่ฝรั่งพันธุ์กลมสีทางลำต้น ให้ผลดีกว่าให้ทางราก” ในการทดลองนี้

1. จะต้องจัดอะไรให้เหมือนกันบ้าง
2. จะต้องจัดอะไรให้ต่างกัน
3. เมื่อจัดสถานการณ์ดังนี้แล้ว สิ่งที่เราจะต้องติดตามดูคืออะไร

1.3.4. ทักษะการทดลอง

จากการศึกษาค้นคว้าได้มีนักการศึกษาได้ให้ความหมายของทักษะการทดลองไว้หลายท่าน ดังนี้

พิศาส สร้อยรุห์ (2529 : 239) กล่าวว่า การทดลองเป็นขั้นตอนที่สำคัญอย่างหนึ่งในการทำงานของนักวิทยาศาสตร์ที่จะช่วยให้สามารถตรวจสอบได้ว่า สมมุติฐานที่ตั้งขึ้นมานั้น ยอมรับได้หรือไม่เพียงใด ซึ่งเป็นการช่วยให้ได้คำตอบที่แท้จริงของปัญหาที่เขาประสบ ถ้าปราศจากการทดลอง อาจทำให้ปัญหาของนักวิทยาศาสตร์บางปัญหา ไม่สามารถจะหาคำตอบที่แท้จริงได้ กล่าวคือ เป็นเพียงสมมุติฐานที่ขาดการตรวจสอบนั่นเอง

สุวัณณ์ นิยมคำ (2531 : 264) กล่าวว่า การทดลองคือการสร้างสถานการณ์จำลองขึ้นมาอย่างหนึ่ง เพื่อจะได้สังเกตผลที่เกิดขึ้นภายใต้การควบคุมเงื่อนไขต่าง ๆ ไว้แล้ว เพื่อจะได้ตรวจสอบซ้ำได้อีก การทดลองมีวัตถุประสงค์ที่จะตรวจสอบสมมุติฐาน รูปแบบหรือโมเดลที่ยังสงสัยอยู่ โดยการกระทำให้วัตถุหรือ

เหตุการณ์เกิดการเปลี่ยนแปลง แล้วนำผลมาเปรียบเทียบกัน การทดลองจะต้องประกอบด้วยกิจกรรม 2 ตอน คือ กิจกรรมภาคออกแบบการทดลองกับกิจกรรมภาคปฏิบัติทดลอง

สมาคมอเมริกันเพื่อความก้าวหน้าทางวิทยาศาสตร์ (AAAS. 1970) ให้ความหมายของทักษะการทดลองว่า "เป็นความสามารถในการจัดกระบวนการปฏิบัติทดลอง เพื่อตรวจสอบสมมุติฐานที่กำหนดไว้"

กูต (Good. 1973 : 227) ให้ความหมายของทักษะการทดลองว่า "เป็นความสามารถในการดำเนินตามแผน โดยมีการควบคุมตัวแปรต่าง ๆ เพื่อทดสอบสมมุติฐานที่ตั้งไว้"

จากแนวคิดและความหมายของทักษะการทดลอง ผู้วิจัยขอสรุปว่า ทักษะการทดลอง หมายถึง ความสามารถจัดกระบวนการปฏิบัติทดลอง เพื่อหาคำตอบ พิสูจน์หรือตรวจสอบสมมุติฐานที่กำหนดไว้ หรือเพื่อยืนยันความจริงในสิ่งที่สงสัย ในสิ่งที่อยากจะรู้คำตอบ การทดลองเป็นการนำเอาทักษะกระบวนการทางวิทยาศาสตร์หลาย ๆ อย่างมาใช้ในการหาคำตอบ และแบ่งประเภทของการทดลองออกเป็น 3 ประเภท คือ

1. การทดลองแบบไม่แบ่งกลุ่มเปรียบเทียบ เป็นการทดลองเพียงกลุ่มเดียวโดยไม่ต้องใช้กลุ่มเปรียบเทียบหรือทำการทดลองเพื่อหาคำตอบจากกลุ่มทดลองนี้ โดยไม่จำเป็นต้องตั้งสมมุติฐาน ลักษณะของการทดลองมีตั้งแต่การทดลองง่าย ๆ ไปจนถึงการทดลองที่มีแบบแผน นั่นคือ มีการทดลองบางอย่างไม่จำเป็นต้องแบ่งกลุ่มเปรียบเทียบ คือ มีกลุ่มทดลองกับกลุ่มควบคุม ก็สามารถสรุปผลการทดลองได้ เช่น การหาจุดเดือดของน้ำ การทดลองทำได้ด้วยการนำ น้ำกลั่นใส่หลอดทดลอง ปิดปากหลอดด้วยจุกยางที่มีเทอร์มอมิเตอร์และหลอดนำก๊าซเสียบอยู่ ต้มน้ำในหลอดทดลอง อ่าน บันทึกอุณหภูมิทุก 1 นาที และสรุปผลการทดลอง หรือการทดลองของอาร์คีเมเดซก็ไม่ได้มีการแบ่งกลุ่มเปรียบเทียบ อาร์คีเมเดซทดลองพบว่า "น้ำหนักของวัตถุที่หายไปเมื่อชั่งในของเหลว จะเท่ากับน้ำหนักของของเหลวที่ถูกวัตถุนั้นแทนที่" เป็นต้น

2. การทดลองแบบแบ่งกลุ่มเปรียบเทียบ เป็นการทดลองโดยการแบ่งออกเป็น 2 กลุ่มเปรียบเทียบผลการทดลองระหว่าง 2 กลุ่มนี้ กลุ่มหนึ่งเรียกว่า "กลุ่มทดลอง" อีกกลุ่มหนึ่งเรียกว่า "กลุ่มควบคุม" ทั้ง 2 กลุ่ม จะกำหนดตัวแปรต้นให้แตกต่างกัน ส่วนตัวแปรที่ต้องควบคุม จะต้องกำหนดหรือจัดให้มีคุณสมบัติและสิ่งแวดล้อมเหมือนกับทุกประการแล้วติดตามดูผลการทดลองของตัวแปรต้น ซึ่งเราเรียกว่า "ตัวแปรตาม" จากสมมุติฐานว่า "หนูที่กินอาหารที่มีโปรตีนจะเจริญเติบโตต่างจากหนูที่กินอาหารที่ไม่มีโปรตีน"

ตัวแปรต้น ได้แก่ อาหารที่มีโปรตีนกับอาหารที่ไม่มีโปรตีน

ตัวแปรตาม ได้แก่ การเจริญเติบโตของหนู

ตัวแปรที่ต้องควบคุม ได้แก่

- 1) ชนิดและขนาดของหนู
- 2) อายุและเพศของหนู
- 3) ปริมาณอาหารและน้ำ
- 4) สถานที่เลี้ยงหนู
- 5) เวลาการให้อาหาร เป็นต้น

3. การทดลองแบบลองผิดลองถูก เป็นการทดลองที่ผู้ทดลองไม่มีหลักเกณฑ์ หรือวิธีทดลองที่แน่นอน ชัดเจนในการหาคำตอบของปัญหา หรือบางปัญหาไม่สามารถทำการทดลองอย่างมีระบบได้ ก็จะใช้วิธีลองผิดลองถูกหรือลองทำ ปรับปรุงแก้ไขหรือเปลี่ยนแปลง จนกว่าจะได้คำตอบหรือมองเห็นวิธีการที่ถูกต้อง การทดลองของนักวิทยาศาสตร์ในสมัยแรก ๆ มักจะเป็นการทดลองแบบลองผิดลองถูก เนื่องจากบางครั้งมีความจำเป็น เพราะบางปัญหาไม่รู้จะตั้งสมมุติฐานอย่างไร ออกแบบการสอนอย่างไร หรือทำการทดลองอย่างมีแบบแผนอย่างไร เป็นต้น หรือการทดลองแบบลองผิดลองถูก ในการค้นพบยาสเตปโตมัยซิน (Streptomycin) ของ ดร.วาคส์แมน (Dr. Waksman) เขาได้นำเอาแบคทีเรียในดิน (Soil Bacteria) มาหลายชนิด แล้วทดลองปล่อยชนิดที่ 1 เข้าไปทำลายเชื้อวัณโรค ถ้าไม่เกิดผลก็ปล่อยชนิดที่ 2, 3, 4,...ตามลำดับ ในที่สุดก็โชคดีพบแบคทีเรียชนิดหนึ่งที่ทำลายเชื้อวัณโรคได้

องค์ประกอบของการทดลอง

วรรณทิพา รอดแรงคำ และพิมพ์พันธ์ เตชะคุปต์ (2532 : VII) ได้จำแนกองค์ประกอบของการทดลองออกเป็น 3 ชั้น คือ

1. การออกแบบการทดลอง หมายถึง การวางแผนการทดลองก่อนลงมือทดสอบจริง
2. การปฏิบัติการทดลอง หมายถึง การลงมือปฏิบัติจริงและใช้อุปกรณ์ได้อย่างถูกต้องและเหมาะสม
3. การบันทึกผลการทดลอง หมายถึง การจดบันทึกข้อมูลที่ได้จากการทดลอง ซึ่งอาจเป็นผลจากการสังเกต การวัดและอื่น ๆ ได้อย่างคล่องแคล่วและถูกต้อง การบันทึกผลการทดลองอาจอยู่ในรูปตารางหรือการเขียนกราฟ ซึ่งโดยทั่วไปจะแสดงค่าของตัวแปรต้นหรือตัวแปรอิสระบนแกนแนวนอนและค่าของตัวแปรตามบนแกนตั้ง โดยเฉพาะในแต่ละแกนต้องใช้สเกลที่เหมาะสม พร้อมทั้งแสดงให้เห็นถึงตำแหน่งของค่าของตัวแปรทั้งสองบนกราฟด้วย

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2536 : 80) จำแนกองค์ประกอบของการทดลองออกเป็น 3 ชั้น คือ

1. การออกแบบการทดลอง หมายถึง การวางแผนการทดลองก่อนลงมือทดลองจริงเพื่อกำหนด
 - 1.1 วิธีการทดลอง (ซึ่งเกี่ยวข้องกับการกำหนดและควบคุมตัวแปร)
 - 1.2 อุปกรณ์และ/หรือสารเคมีที่จะต้องใช้ในการทดลอง
2. การปฏิบัติการทดลอง หมายถึง การลงมือปฏิบัติการทดลองจริง ๆ
3. การบันทึกผลการทดลอง หมายถึง การจดบันทึกข้อมูลที่ได้จากการทดลอง ซึ่งอาจเป็นผลจากการสังเกต การวัดและอื่น ๆ

จากการจำแนกองค์ประกอบของการทดลองดังกล่าว ผู้วิจัยขอจำแนกองค์ประกอบของการทดลองออกเป็น 3 ชั้น เช่นกัน คือ

1. การออกแบบการทดลอง หมายถึง วิธีการวางแผนการทดลองก่อนลงมือปฏิบัติจริง การออกแบบการทดลองจะต้องสอดคล้องกับสมมุติฐานที่ตั้งไว้ และครอบคลุมถึงตัวแปรที่ต้องควบคุม รวมถึงการเลือกใช้อุปกรณ์ เครื่องมือ เครื่องใช้ที่เหมาะสมด้วย การออกแบบการทดลองจะประกอบด้วยส่วนสำคัญ 2 ประการ คือ

1.1 วิธีการทดลอง เป็นขั้นตอนการทดลองตั้งแต่เริ่มต้นจนครบตามขั้นตอน จะเกี่ยวข้องกับการกำหนดและควบคุมตัวแปร

1.2 วัสดุอุปกรณ์ที่ใช้ในการทดลอง

2. การปฏิบัติการทดลอง หมายถึง การลงมือปฏิบัติการทดลองจริง จะดำเนินไปตามขั้นตอนและการใช้วัสดุอุปกรณ์อย่างถูกต้องและเหมาะสม

3. การบันทึกผลการทดลอง หมายถึง การจดบันทึกข้อมูลที่ได้จากการทดลอง ซึ่งอาจเป็นผลมาจากการสังเกต การวัดและอื่น ๆ ได้อย่างคล่องแคล่วและถูกต้อง เช่น การบันทึกข้อมูลในรูปตารางหรือจัดกระทำในรูปของกราฟ เป็นต้น

พฤติกรรมของทักษะการทดลอง

สถาบันการส่งเสริมวิทยาศาสตร์และเทคโนโลยี (2536 : 80) กำหนดพฤติกรรมหรือความสามารถที่แสดงว่า เกิดทักษะการทดลอง คือ

1. ออกแบบการทดลองโดย

1.1 กำหนดวิธีการทดลองได้ถูกต้องและเหมาะสม โดยคำนึงถึงตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุม

1.2 ระบุอุปกรณ์และ/หรือสารเคมีที่จะต้องใช้ในการทดลองได้

2. ปฏิบัติการทดลองและใช้อุปกรณ์ได้อย่างถูกต้องและเหมาะสม

3. บันทึกผลการทดลองได้คล่องแคล่วและถูกต้อง

คำถามที่นำไปสู่ทักษะการทดลอง เช่น

1. จากสมมุติฐานว่า "มุมของปลายรางมีผลต่อระยะที่ตกของลูกปืน"

1.1 นักเรียนจะมีวิธีการทดลองเพื่อตรวจสอบสมมุติฐานข้อนี้ได้อย่างไร

1.2 การทดลองนี้ต้องใช้อุปกรณ์อะไรบ้าง

2. การอ่านอุณหภูมิในของเหลวทำอย่างไร

3. ขณะเผาสารในหลอดทดลองจะต้องทำอย่างไรบ้าง

จากพฤติกรรมหรือความสามารถดังกล่าว ผู้วิจัยขอกำหนดความสามารถของผู้ที่แสดงว่าเกิดทักษะการทดลอง คือ

1. ออกแบบการทดลอง โดยกำหนดวิธีทดลองและอุปกรณ์ได้อย่างถูกต้อง

2. ปฏิบัติทดลองและใช้อุปกรณ์ได้อย่างถูกต้อง

3. บันทึกผลการทดลองได้อย่างถูกต้อง

คำถามที่นำไปสู่ทักษะการทดลอง เช่น

1. จากสมมุติฐานที่ว่า "ระยะที่ปล่อยลูกบอลสู่พื้นดินมากขึ้น ลูกบอลยิ่งกระดอนจากพื้นดินสูงขึ้น"

1.1 นักเรียนคิดว่าจะทำการทดลองเพื่อพิสูจน์สมมุติฐานข้อนี้ได้อย่างไร

1.2 ในการทดลองนี้ต้องใช้อุปกรณ์อะไรบ้าง

2. ต้องใช้ไม้หนีบจับหลอดทดลองตรงไหน

3. ขณะเผาสารในหลอดทดลองจะต้องทำอย่างไร

4. ตักสารเคมีออกจากขวดแล้วต้องทำอย่างไร

1.3.5. ทักษะการตีความหมายข้อมูลและลงข้อสรุป

จากการศึกษาค้นคว้าได้มีนักการศึกษาได้ให้ความหมายของทักษะการตีความหมาย ข้อมูลและลงข้อสรุปไว้หลายท่าน ดังนี้

พิศาล สร้อยธุระ (2529 : 211) กล่าวว่า การนำข้อมูลไปใช้ในการอ้างอิง ในการอธิบายหรือเพื่อจุดประสงค์ใดก็ตามไม่ว่าจะนำไปใช้ทั้งหมดหรือเพียงบางส่วน ผู้นำข้อมูลไปใช้จำเป็นต้องตีความหมายข้อมูลดังกล่าวให้อยู่ในรูปภาษาพูด หรือภาษาเขียนที่สามารถสื่อความหมายกับคนทั่วไปได้ โดยเป็นที่เข้าใจตรงกัน การกระทำเช่นนี้เรียกว่า การตีความหมายข้อมูล

สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี (2534 : 58 – 59) ให้ความหมายของทักษะการตีความหมายข้อมูลและลงข้อสรุปว่า

ทักษะการตีความหมายข้อมูล หมายถึง ความสามารถในการแปลความหมายหรือการบรรยายลักษณะและสมบัติของข้อมูลที่มีอยู่ การตีความหมายข้อมูลในบางครั้งอาจต้องใช้ทักษะอื่น ๆ ด้วย เช่น ทักษะการสังเกต ทักษะการคำนวณ เป็นต้น

การลงข้อสรุป หมายถึง การสรุปความสัมพันธ์ของข้อมูลทั้งหมด

สมาคมอเมริกันเพื่อความก้าวหน้าทางวิทยาศาสตร์ (AAAS. 1970) ได้ให้ความหมายของทักษะการตีความหมายข้อมูลและลงข้อสรุป หมายถึง ความสามารถในการแปลความหมายหรือการบรรยายลักษณะและสมบัติของข้อมูลที่มีอยู่

กูต (Good. 1973 : 313, 125) กล่าวว่า ทักษะการตีความหมายข้อมูล หมายถึง ความสามารถในการแปลความหมายหรืออธิบายลักษณะของข้อมูล ส่วนทักษะการลงข้อสรุป หมายถึง ความสามารถในการสรุปความสัมพันธ์ของข้อมูลทั้งหมด หรือการตัดสินใจในขั้นสุดท้าย

จากแนวคิดและความหมายของทักษะการตีความหมายข้อมูลและลงข้อสรุป ผู้วิจัยจึงขอกำหนดความหมายและตัวอย่างประกอบคือ

✓ ทักษะการตีความหมายข้อมูล หมายถึง ความสามารถในการแปลความหรือบรรยายความหมายของข้อมูลที่ได้จัดกระทำแล้ว หรือข้อมูลที่มีอยู่ในรูปแบบที่ใช้ในการสื่อความหมายแล้ว การตีความหมายข้อมูลบางครั้งอาจต้องใช้ทักษะกระบวนการวิทยาศาสตร์อื่น ๆ เช่น ทักษะการสังเกต ทักษะการวัด ทักษะการคำนวณ ทักษะการจัดกระทำและสื่อความหมาย ข้อมูล เป็นต้น

✓ การลงข้อสรุป หมายถึง การสรุปความสัมพันธ์ของข้อมูล ที่เกี่ยวข้องกับความสัมพันธ์ของตัวแปรต่าง ๆ ที่ต้องการ การลงข้อสรุปจะเกี่ยวข้องกับสมมุติฐานที่ตั้งไว้ก่อนการทดลอง ✓

ทักษะการตีความหมายข้อมูลแตกต่างจากทักษะการพยากรณ์ กล่าวคือ ทักษะการตีความหมายข้อมูลเป็นความสามารถในการตีความ แปลความหมายหรือบรรยายความหมายของข้อมูลที่ได้จัดกระทำแล้ว หรือข้อมูลที่มีอยู่ ส่วนทักษะการพยากรณ์ เป็นความสามารถในการทำนายหรือคาดคะเนคำตอบล่วงหน้าโดยอาศัยข้อมูลที่มีอยู่ ซึ่งอาจจะพยากรณ์ภายในหรือนอกขอบเขตของข้อมูล

สำหรับทักษะการลงข้อสรุปจะต่างจากทักษะการลงความคิดเห็นจากข้อมูล กล่าวคือ/ทักษะการลงข้อสรุป เป็นความสามารถในการสรุปความสัมพันธ์ของข้อมูล ซึ่งจะเกี่ยวกับความสัมพันธ์ของตัวแปรต่าง ๆ และเป็นการสรุปผลการทดลองเพื่อจัดตั้งเป็นหลักการ กฎหรือทฤษฎีต่อไป ส่วนทักษะการลงความคิดเห็นจากข้อมูล คือความสามารถในการเพิ่มความคิดเห็นหรือการอธิบายผลที่ได้จากการสังเกต การวัดหรือการทดลองข้อมูล โดยอาศัยความรู้และประสบการณ์เดิม ซึ่งการลงความคิดเห็นจากข้อมูลจะเป็นสิ่งที่ยังไม่มั่นใจ อาจจะถูกหรือผิดก็ได้

ประเภทหรือลักษณะของการลงข้อสรุป

หลังจากตีความหมายข้อมูลแล้ว ขั้นตอนต่อไปก็คือการลงข้อสรุป วิธีการลงข้อสรุปที่ง่าย ที่สุด คือการมองย้อนไปดูปัญหาหรือสมมุติฐานก่อน แล้วจึงพิจารณาว่า เมื่อรวบรวมเอาความหมายจากข้อมูลที่ได้เข้าด้วยกันแล้ว ตอบรับหรือปฏิเสธสมมุติฐานนั้น อย่างไรก็ตามข้อสรุปที่ได้จะต้องครอบคลุมความหมายที่ดีความได้จากข้อมูลทั้งหมด ข้อความที่ดีความหมายได้ทุกประเด็นจะต้องไม่ขัดแย้งกับการลงข้อสรุป

สมชัย โกมล และคณะ (2525 : 319) ได้กล่าวถึงลักษณะการลงข้อสรุปไว้ 4 ลักษณะ คือ

1. ลงข้อสรุป เพื่อหาความสัมพันธ์ระหว่างองค์ประกอบที่มีอยู่ในปรากฏการณ์นั้น ข้อสรุปที่ได้แบบนี้บอกเฉพาะความสัมพันธ์ระหว่างตัวแปร แต่ไม่อาจบอกความเป็นเหตุเป็นผลต่อกัน
2. ลงข้อสรุป เพื่อหาความสัมพันธ์ขององค์ประกอบที่มีอยู่ในปรากฏการณ์นั้น ในเชิงที่เป็นเหตุและเป็นผลซึ่งกันและกัน ข้อสรุปที่ได้แบบนี้จะบอกว่าอะไร เป็นสาเหตุของอะไรได้
3. ลงข้อสรุป เพื่อหาความสัมพันธ์ในลักษณะที่เป็นแนวโน้มหรือระบบการทำงาน
4. ลงข้อสรุป เพื่อหาหน้าที่เฉพาะ สมบัติเฉพาะหรือจัดประเภทของสิ่งที่ศึกษา

แอนเดอร์สัน (Anderson. 1976 : 43) ได้จำแนกประเภทของการลงข้อสรุปไว้ 5 ประเภท คือ

1. การลงข้อสรุปรวมทั่วไป เพื่อหาหน้าที่เฉพาะของหน่วยใดหน่วยหนึ่ง ในระบบอันหนึ่ง
2. การลงข้อสรุปรวมทั่วไป เพื่อหาการทำงานของระบบและรูปลักษณะของระบบอันหนึ่ง
3. การลงข้อสรุปรวมทั่วไป เพื่อหาการจัดประเภทใหม่ของสิ่งต่าง ๆ
4. การลงข้อสรุปรวมทั่วไป เพื่อหาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ
5. การลงข้อสรุปรวมทั่วไป เพื่อหาความสัมพันธ์ที่เป็นเหตุเป็นผลต่อกัน

พฤติกรรมของทักษะการตีความหมายข้อมูลและลงข้อสรุป

สวบันการสอนวิทยาศาสตร์และเทคโนโลยี (2534 : 59) ได้กำหนดพฤติกรรมหรือความสามารถผู้ที่แสดงว่า เกิดของทักษะการตีความหมายข้อมูลและลงข้อสรุป คือ

1. แปลความหมายหรือบรรยายลักษณะและสมมติของข้อมูลที่มีอยู่ได้
2. บอกความสัมพันธ์ของข้อมูลที่มีอยู่ได้

คำถามที่นำไปสู่ทักษะการตีความหมายข้อมูลและลงข้อสรุป เช่น

1. สมมติว่าทดลองกิจกรรมชุดกลิ้งหอยและได้บันทึกผลการทดลองดังนี้

ตาราง 2 ผลการทดลองกิจกรรมชุดกลิ้งหูกคะเมน

ลูกกลิ้งลูกที่	ลักษณะของลูกกลิ้ง	ผลการทดลอง
1	มีลูกแก้ว	กลิ้งไปตามราง
2	มีลูกแก้วยึดติดกับฐาน	ไม่กลิ้งไปตามราง
3	ไม่มีลูกแก้ว	ถ้าวางตั้งจะกลิ้ง แต่ถ้านอนจะไม่กลิ้ง
4	มีลูกแก้ว	กลิ้งไปตามรางแต่วิ่ง

จะสรุปข้อมูลเกี่ยวกับการกลิ้งของลูกกลิ้งได้อย่างไร

2. จากข้อมูลลักษณะของลูกกลิ้งมีความสัมพันธ์กับการเคลื่อนที่อย่างไร

จากพฤติกรรมหรือความสามารถดังกล่าว ผู้วิจัยขอกำหนดความสามารถของผู้ที่แสดงว่าเกิดของทักษะการตีความหมายข้อมูลและลงข้อสรุป คือ

1. แปลความหมาย หรือบรรยายลักษณะและสมบัติของข้อมูลที่มีอยู่ได้
2. บอกความสัมพันธ์ของข้อมูลที่มีอยู่และลงข้อสรุปได้

คำถามที่นำไปสู่ของทักษะการตีความหมายข้อมูลและลงข้อสรุป เช่น

1. จากการวัดความดันบรรยากาศและความหนาแน่นของอากาศที่ระดับความสูงต่าง ๆ ได้ผลการทดลองดังนี้

ตาราง 3 ผลการวัดความดันบรรยากาศและความหนาแน่นของอากาศที่ระดับความสูงต่าง ๆ

ความสูงจากระดับน้ำทะเล (km)	ความดันบรรยากาศ (mm ของปรอท)	ความหนาแน่นของอากาศ (g/cm ³)
0	760	0.00130
2	600	0.00100
4	470	0.00082
6	360	0.00066
8	280	0.00052
10	210	0.00041
12	160	0.00031

1.1 ความดันบรรยากาศที่ระดับความสูง 5 กิโลเมตร จากระดับน้ำทะเลมีค่าเท่าใด

1.2 ที่ระดับความสูงขึ้นไป ความดันบรรยากาศและความหนาแน่นของอากาศจะเป็นอย่างไร

2. จากตารางข้อมูลข้างต้น สรุปความสัมพันธ์ระหว่างระดับความสูง ความดันบรรยากาศและความหนาแน่นของอากาศได้อย่างไร

จากทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการทั้ง 5 ทักษะนี้ ผู้วิจัยจึงได้นำเอาพฤติกรรมที่ได้แสดงว่าเกิดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการในด้านต่าง ๆ มาเป็นแนวทางการออกแบบสร้างเป็นแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ ชนิดที่มีสถานการณ์กำหนดให้ตอบและอธิบายประกอบ พร้อมทั้งนำไปสร้างกฎเกณฑ์การให้คะแนน (Scoring Rubric) และหาความเชื่อมั่นของเกณฑ์การให้คะแนนในกรณีที่มีรูปแบบการตรวจให้คะแนนที่แตกต่างกันต่อไป

1.4 การประเมินผลทักษะกระบวนการทางวิทยาศาสตร์

ซันด์ และ พิคาร์ด (Sund and Picard. 1972 : 31-25) กล่าวถึง การประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ว่า ต้องศึกษาวัตถุประสงค์เชิงพฤติกรรมของแต่ละทักษะ เพื่อใช้เป็นแนวทางการประเมินผลดูว่า นักเรียนมีความสามารถในการใช้ทักษะกระบวนการทางวิทยาศาสตร์ในด้านต่าง ๆ หรือไม่ มากน้อยเพียงใด

หน่วยทดสอบและประเมินผลสถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี ได้เสนอแนะแนวทางในการสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ไว้ดังนี้ (สสวท. 2518 : 5)

1. กำหนดความมุ่งหมายเชิงพฤติกรรม ซึ่งจะต้องแจกแจงให้ชัดเจน โดยครูต้องศึกษาจุดมุ่งหมายในแต่ละทักษะให้เข้าใจ แล้วมาแจกแจงให้เป็นจุดมุ่งหมายเชิงพฤติกรรม ซึ่งจะมีทั้งภาคสถานการณ์ ภาคพฤติกรรมที่คาดหวัง และเกณฑ์ในการกำหนดพฤติกรรมนั้น ๆ
2. การเลือกเนื้อหาที่จะวัด หมายถึง การเลือกความมุ่งหมายเชิงพฤติกรรมกับเนื้อหาที่จำเป็นที่ขาดเสียมิได้ในบทหนึ่ง ๆ ควรจะกำหนดว่าทักษะใด เนื้อหาใดเป็นสิ่งที่ขาดมิได้ ทักษะนั้นและเนื้อหานั้นก็ควรจะปรากฏในข้อสอบ
3. การสร้างตารางเพื่อกำหนดเนื้อหาและพฤติกรรมทักษะ ซึ่งมีความมุ่งหมายอยู่ที่จะกำหนดว่าจะวัดทักษะหรือพฤติกรรมได้เท่าไร อย่างละกี่ข้อ จะได้ไม่บกพร่อง นอกจากนั้นผู้ออกข้อสอบยังจะต้องทราบต่อไปอีกว่า ข้อสอบวัดพฤติกรรมทักษะใดมีสัดส่วนมากนักน้อยเพียงใด
4. การเลือกแนวทางการออกข้อสอบ ควรจะถือหลักว่าจะใช้การสอบแบบใด จึงจะวัดพฤติกรรมนั้นได้ตรงและถูกต้องเหมาะสมที่สุด ตลอดจนทั้งเหมาะสมกับวัยของเด็ก ประหยัดเวลาและง่ายต่อการปฏิบัติด้วย

นอกจากนี้ ยังได้เสนอลักษณะข้อทดสอบเพื่อวัดความสามารถในการดำเนินการตามกระบวนการทางวิทยาศาสตร์ ดังนี้ (สสวท. 2519 : 10)

1. การสร้างสถานการณ์
 - 1.1 สถานการณ์ที่สร้างขึ้น จะเป็นสถานการณ์สมมติ หรือนำมาจากเอกสารอื่นใดก็ตาม จะต้องมีความยากง่ายเหมาะสมกับระดับชั้นของนักเรียน
 - 1.2 ใช้คำพูดที่เข้าใจง่าย ศัพท์เทคนิคต้องไม่นอกเหนือจากที่นักเรียนเรียนรู้มาแล้ว
 - 1.3 สถานการณ์นั้น ๆ ต้องมีความเป็นไปได้ จะต้องเป็นจริงสมเหตุผล
 - 1.4 ถ้าเป็นเรื่องที่มีหน่วยวัด จะต้องระบุให้ชัดเจนว่าเป็นหน่วยใด
 - 1.5 สถานการณ์ที่ยกมาต้องสั้น กระชับ อ่านเข้าใจได้ง่าย แต่ละสถานการณ์ควรใช้สำหรับถามได้มากกว่า 1 ข้อ เพื่อมิให้นักเรียนเสียเวลาในการอ่านมากเกินไปจนความจำเป็น

2. การสร้างคำถาม คำถามที่จะให้ตอบตามสถานการณ์ที่ยกมาจะมีคุณสมบัติดังนี้

2.1 ถามในเรื่องที่ต้องใช้ความสามารถในด้านกระบวนการทางวิทยาศาสตร์ ไม่ถามเรื่องที่เป็นความรู้-ความจำ

2.2 ไม่ถามถึงปัญหา หรือสมมุติฐานที่เคยอภิปราย หรือสรุปกันมาแล้วเพราะจะกลายเป็นความจำทั้ง ๆ ที่คำถามเหมือนกับจะวัดทักษะกระบวนการทางวิทยาศาสตร์

2.3 ใช้คำถามรัดกุม บังคับว่าจะให้ตอบเรื่องใด แม้ว่าบางคำถามจะให้เห็นความคิดเห็นได้แตกต่างกัน แต่ก็ต้องเป็นความคิดเห็นเกี่ยวกับเรื่องนั้น ๆ โดยเฉพาะ

2.4 ข้อความที่จะให้ตอบแต่ละคำถาม ควรเป็นตอนละเรื่อง และกำหนดระดับคะแนนให้เหมาะสม ถ้าเป็นไปได้ควรให้คะแนนเป็น 1 ถ้าตอบถูก และให้ 0 ถ้าตอบผิด

3. การตรวจให้คะแนน

ถ้าเป็นข้อทดสอบให้ตอบสั้น ๆ แม้จะตั้งคำถามที่ผู้ถามคิดว่าจำเพาะเจาะจง คำตอบน่าจะแน่นอน แต่ในการตรวจจะต้องดูที่เหตุผลของนักเรียนบางคนที่ตอบแตกต่างไปจากเกณฑ์ที่ตั้งไว้ด้วย ถ้าเหตุผลถูกต้องก็ต้องยอมรับ

จากแนวคิดการประเมินผลทักษะกระบวนการทางวิทยาศาสตร์ดังกล่าว ผู้วิจัยจึงได้ใช้เป็นแนวทางในการสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ โดยกำหนดสถานการณ์ให้ตอบคำตอบของผู้เรียนจะบ่งชี้ถึงความสามารถด้านทักษะกระบวนการในแต่ละทักษะ

2. แบบทดสอบความเรียง (Essay Type Test)

2.1 ความหมายของแบบทดสอบความเรียง

กลอนลันด์ (Gronlund. 1977 : 75) สรุปความหมายของแบบทดสอบความเรียงไว้ว่า ถึงแม้จะกล่าวว่าแบบทดสอบแบบความเรียงจะไม่มีประสิทธิภาพในการวัดผลการเรียนรู้ แต่ข้อสอบแบบนี้จะให้อิสระในการตอบ ซึ่งเป็นสิ่งที่จำเป็นในการวัดผลที่ซับซ้อน รวมไปถึงการวัดความสามารถในการสร้าง การจัดระเบียบ การรวมเป็นหนึ่งเดียว การแสดงออกและพฤติกรรมที่คล้ายคลึงกันซึ่งก่อให้เกิดการสร้างและการสังเคราะห์ความคิด

เวียสม่า และ เจอร์ส (Wiersma and Jurs. 1985 : 165) ได้สรุปความหมายของแบบทดสอบความเรียงว่าเป็นแบบทดสอบที่ให้นักเรียนแสดงความต้องการที่จะตอบอย่างมีโครงสร้าง ใช้ข้อมูลที่ละเอียดและยาวมากขึ้น จะประกอบด้วยรายการบรรยายหลาย ๆ ส่วน

ฮอบกินส์ และ แอนเตส (Hopkins & Antes. 1990 : 231) ให้ความหมายของแบบทดสอบความเรียง (Essay Items) ไว้ว่า เป็นข้อสอบประเภทขยายความ (Constructed – response Items) ที่สมบูรณ์ที่สุดเป็นข้อสอบที่ต้องการเห็นพัฒนาการในการตอบคำถามของนักเรียน

กายย์ (Gay, L.R., 1992 : 224) กล่าวว่า แบบทดสอบความเรียง เป็นแบบทดสอบที่ประกอบด้วย หัวข้อที่ต้องการให้นักเรียนเขียนตอบ โดยปกติแล้วจะเป็นคำตอบที่ยาวจริง ๆ แล้ว แบบทดสอบความเรียงไม่เหมาะสมสำหรับวัดผลระดับความรู้ความจำ แต่เหมาะสำหรับวัดผลในระดับสูง เพราะแบบทดสอบแบบนี้จะต้องมีการสังเคราะห์ และการประเมินค่า

เวอร์ต (Ward, 1999 : 129) ให้ความหมายของแบบทดสอบความเรียงว่า เป็นแบบทดสอบที่วัดทักษะขั้นสูง ยากที่จะสร้างด้วยแบบทดสอบแบบปรนัยได้ แบบทดสอบประเภทนี้ต้องการให้นักเรียนแสดงความรู้ กระบวนการของความรู้ การรวบรวมข้อมูล และนำเสนอความรู้ในรูปของเหตุผลที่ประติดประต่อกันได้

ลวัน สายยศ และอังคณา สายยศ (2539 : 86) ได้กล่าวถึง แบบทดสอบความเรียงว่ามีจุดประสงค์วัดความสามารถในการบรรยาย อธิบายและแสดงเหตุผลตามความคิดเห็นของตน อาจจำกัดความยาวหรือให้เขียนได้ตามสบายก็ได้

โกวิท ประวาลพุกษ์ และ สมศักดิ์ สินธุระเวชญ์ (2525 : 101) ให้นิยามแบบทดสอบความเรียงว่าเป็นข้อสอบที่ให้ได้มีอิสระในการตอบเต็มที่ ให้โอกาสเด็กได้แสดงว่าเขารู้อะไร สามารถวัดสมรรถภาพด้านบูรณาการความคิด ความสามารถในการจัดรวบรวมความคิดที่สมบูรณ์ (Organization Ability) การสังเคราะห์ และการประเมินค่าอย่างมีประสิทธิภาพ เช่น ความรู้ในการแก้ปัญหาใหม่ ๆ และเป็นต้นคิดหรือค้นหาวิธีการใหม่ ๆ เพื่อแก้ปัญหาต่าง ๆ ได้

วิญญา วิศาลาภรณ์ (2533 : 60) ให้ความหมายของแบบทดสอบความเรียง (Essay Test) หรือ (Subjective Test) เป็นข้อสอบที่ให้เสรีภาพเป็นอย่างมากแก่นักเรียนในการตอบ นักเรียนมีอิสระในการที่จะเสนอความคิดด้วยภาษาของตนเอง นักเรียนต้องใช้ความสามารถอธิบายปัญหาต่าง ๆ ได้ ดังนั้น ข้อสอบแบบความเรียงหรือข้อสอบอัตนัยจึงสามารถใช้วัดความรู้ความคิดขั้นสูงที่ซับซ้อนได้ วัดทัศนคติได้ นักเรียนต้องมีความสามารถในการจัดระเบียบความรู้ความคิด แสดงความรู้เริ่มและรู้จักสังเคราะห์

พวงรัตน์ ทวีรัตน์ (2530 : 110) ให้ความหมายของแบบทดสอบอัตนัย หรือ แบบทดสอบความเรียง (Essay Test) ว่าเป็นแบบทดสอบที่มีคุณค่ามากชนิดหนึ่ง สามารถใช้วัดพฤติกรรมด้านพุทธิพิสัยหรือด้านความรู้และความคิดระดับสูงได้ดี เช่น วัดพฤติกรรมขั้นการสังเคราะห์ และการประเมินค่า เป็นต้น

รุจิร ภูสาระ (2520 : 31) กล่าวว่า การสอบแบบบรรยาย (Essay Examination) เป็นการตอบคำถามซึ่งมีจำนวนข้อจำกัด โดยปกติมีประมาณ 5 – 10 ข้อ ครูเป็นผู้ตั้งคำถามให้นักเรียนเป็นผู้ตอบ ครูจะถามในสิ่งที่นักเรียนได้เรียนมา เพื่อให้นักเรียนตอบตามความสามารถของแต่ละคน และมักตั้งคำถามว่า ใคร เมื่อไร อะไร หรือ ทำไม เพื่อถามนักเรียนถึงชื่อ สถานที่ ให้อภิปราย ประเมินค่า หาเหตุผล หาความหมายของสถานการณ์ต่าง ๆ เป็นต้น

บุญชม ศรีสะอาด (2534 : 80) กล่าวถึงข้อสอบอัตนัย (Essay) ว่า เป็นข้อสอบที่ให้ผู้สอบตอบ โดยการเขียนตามความคิดของตนเอง แม้ว่าทั่วไปจะได้รับการนำไปใช้น้อย แต่ก็ เป็นเครื่องมือที่จำเป็นในการ วัดผลโดยเฉพาะในการวัดเกี่ยวกับความคิดเห็น เจตคติ ความสามารถในการคิดสังเคราะห์ (Synthesis) ความสามารถในการอธิบายให้คนอื่นเข้าใจ ความสามารถในการบูรณาการความรู้ การวัดในลักษณะเหล่านี้ ข้อสอบแบบอัตนัยจะวัดได้ดีมาก

จากการศึกษานิยามของแบบทดสอบความเรียงจากนักการศึกษาหลาย ๆ ท่านที่กล่าวมานี้ ผู้วิจัย สรุปได้ว่า แบบทดสอบความเรียงหรือเรียกโดยทั่วไปว่าแบบทดสอบอัตนัย เป็นแบบทดสอบที่ใช้เพื่อการวัด ผลสัมฤทธิ์ทางการเรียนในพฤติกรรมการเรียนรู้ชั้นสูง ที่ต้องการให้นักเรียนแสดงความคิดของตนเองออกมา ด้วยภาษาและถ้อยคำที่ต้องอาศัยการเรียบเรียง และรวบรวมความรู้มาตอบด้วยตนเอง

2.2 การสร้างแบบทดสอบความเรียง

ได้มีนักวัดผลการศึกษาทั้งในและต่างประเทศนำเสนอขั้นตอนในการสร้างแบบทดสอบความเรียงไว้ หลาย ๆ ท่าน ดังนี้

มาแชล และ เฮลส์ (Marshall and Hales, 1971 : 31-32) เสนอแนะแนวทางการวางแผนการเขียน ข้อสอบความเรียงไว้ดังนี้

1. ให้มีเวลาพอเพียงในการสร้างข้อสอบแบบความเรียง การเขียนข้อสอบที่ดีนั้นต้องการเวลา ก่อน จะนำข้อสอบไปใช้ควรจะมีแน่ใจว่าข้อสอบนั้นมีคำสั่ง และคำถามชัดเจนไม่คลุมเครือ มีขอบเขตของการตอบ และเป็นข้อสอบที่วัดสิ่งที่ครูต้องการวัด
2. ข้อสอบแบบความเรียง จะประกอบด้วยปัญหาที่ต้องการจะให้แก้ไข เนื่องจากข้อสอบความเรียง เหมาะที่จะวัดพฤติกรรมสมรรถนะชั้นสูง และเป็นกระบวนการที่เกี่ยวข้องกับการแก้ปัญหา
3. ควรจะนิยามปัญหาให้กระจ่างชัด ข้อสอบแบบความเรียงจะขาดความเที่ยงตรง ถ้าหากนักเรียน ไม่สามารถตีความหมายของปัญหา นักเรียนทุกคนจะต้องตีความปัญหาให้ออกมาเหมือนกัน เพราะนักเรียน จะต้องตอบ “ข้อสอบเดียวกัน”
4. ควรจะมีขอบเขตของปัญหา เป็นการยากที่นักเรียนจะตอบปัญหาที่คำถามคลุมเครือมากเกินไป การไม่จำกัดขอบเขตปัญหา จะทำให้นักเรียนตอบอะไรก็ได้ที่เขารู้และเกี่ยวข้องกับปัญหานั้น ปัญหาที่ไม่มี ขอบเขตจะยั่วให้นักเรียนเดาและทำให้การทดสอบมีความเที่ยงตรงต่ำ
5. ควรระบุคำแนะนำในการตอบข้อสอบให้ชัดเจน นักเรียนจะต้องเข้าใจตรงกันว่าข้อสอบแต่ละข้อ ครอบคลุมเท่าไร ควรจะตอบข้อสอบแต่ละข้อโดยใช้เวลาเท่าไร ควรจะใช้ข้อมูลอะไรในการตอบ รูปแบบในการ เขียนตอบ ควรจะเป็นอย่างไร ควรจะหลีกเลี่ยงคำถามที่นักเรียนต้องเดาว่าครูต้องการอะไร
6. ไม่ควรมีข้อเลือกตอบ หากมีข้อเลือกตอบที่มีความสามารถต่างกัน ก็จะเลือกตอบข้อคำถามที่ ต่างกัน ตัวครูเองก็มีความโน้มเอียงในการให้คะแนนบางข้อมากกว่าบางข้อ จึงเป็นการยากที่จะเปรียบเทียบ ความสามารถของนักเรียนแต่ละคน ยิ่งไปกว่านั้น นักเรียนอาจเลือกดูหนังสือเพียงบางส่วนเพื่อการสอบก็ได้
7. ควรจะทำคำถามไว้ล่วงหน้า หลังจากสร้างคำถามแล้ว ครูควรจะเขียนคำตอบที่ตนเองเห็นว่า เหมาะสมที่สุด วิธีการนี้จะทำให้ครูสามารถพบข้อบกพร่องของคำถาม เช่น ความคลุมเครือ ความไม่รัดกุม ของคำถาม เป็นต้น

กลอนลันด์ (Gronlund, 1977 : 81-83) เสนอกฎการตั้งคำถามแบบทดสอบความเรียง เพื่อให้เกิดคุณภาพของข้อสอบมีคุณภาพสูงไว้ดังนี้

1. ให้ใช้แบบทดสอบความเรียงเพื่อวัดผลการเรียนรู้ที่ซับซ้อนเท่านั้น
2. เชื่อมโยงคำถามกับผลการเรียนรู้ที่ต้องการวัดโดยตรงที่สุดเท่าที่จะเป็นไปได้
3. สร้างคำถามที่นักเรียนเข้าใจได้อย่างชัดเจน
4. ไม่มีตัวเลือกของคำถามให้นักเรียนเลือกตอบ ยกเว้นว่าผลการเรียนรู้ต้องการให้เป็นไปเช่นนั้น
5. ให้อเวลาในการตอบอย่างเพียงพอ และระบุเวลาไว้ในแต่ละคำถามด้วย

บุญเชิด ภิญโญนนตพงษ์ (2525 :38-41) ได้ให้ข้อแนะนำในการสร้างข้อสอบอัตนัยให้มีคุณภาพ ต้องคำนึงถึงจุดอ่อน 2 ประการ คือ ทำอย่างไรจึงจะสร้างคำถามให้สามารถวัดในสิ่งที่ต้องการวัดได้ และ ทำอย่างไรจึงจะทำให้การตรวจข้อสอบอัตนัยมีความเชื่อถือได้มากที่สุด ดังนั้น ในการสร้างข้อสอบจะต้องเขียนข้อคำถาม และกำหนดแนวการตอบและเกณฑ์สำหรับตัดสินให้คะแนนอย่างละเอียดด้วย สาเหตุหนึ่งที่ทำให้คะแนนในการตรวจเชื่อถือไม่ค่อยได้ เพราะการเฉลยส่วนนี้ไป ดังนั้น การสร้างข้อสอบอัตนัยจึงควรคำนึงถึงข้อแนะนำต่อไปนี้

1. วิเคราะห์จุดประสงค์ของวิชาที่จะออกข้อสอบให้ละเอียดครอบคลุม และเลือกรูปแบบคำถามให้เหมาะสมกับจุดประสงค์นั้น ๆ
2. เขียนคำถามให้ชัดเจนและตรงจุด ระบุให้แน่ชัดว่าต้องการให้ผู้ตอบทำอย่างไร
3. เขียนคำชี้แจงให้ละเอียดและชัดเจน เช่น
คำชี้แจง ข้อสอบมีทั้งหมด 5 ข้อ แต่ละข้อให้คะแนน 20 คะแนน ผู้ตอบต้องแบ่งเวลาในการตอบแต่ละข้อให้เท่า ๆ กัน การให้คะแนนจะพิจารณาถึงความสามารถในการเรียบเรียงความคิด ความถูกต้องของเนื้อหา ตัวสะกดและหลักไวยากรณ์ และความสะอาดเรียบร้อย
4. ควรเขียนข้อสอบแบบจำกัดคำตอบ มากกว่าข้อสอบแบบไม่จำกัดคำตอบ เพื่อให้การตรวจมีความเชื่อมั่นสูงขึ้น
5. หลีกเลี่ยงคำถามประเภทวัดความจำอย่างเดียว เพราะการวัดความจำสามารถใช้ข้อสอบปรนัยอื่น ๆ วัดได้ดีกว่า
6. ควรเขียนข้อสอบให้มากข้อ และหลายรูปแบบ เพื่อให้ครอบคลุมเนื้อหา และสิ่งที่ต้องการจะวัด ซึ่งช่วยให้การวัดมีความเที่ยงตรงตามเนื้อหา (Content Validity) สูงขึ้น
7. ควรกำหนดเวลาที่ใช้ตอบแต่ละข้ออย่างคร่าว ๆ ดังในข้อ 3 การคาดคะเนเวลาที่จะใช้เขียนตอบจะเป็นการช่วยให้ผู้ตอบสามารถกำหนดขอบเขตในการตอบได้เหมาะสม และสามารถทำข้ออื่น ๆ ได้ทัน
8. ควรเฉลยคำตอบไว้ทั้งคำตอบที่ถูกต้องและแนวคำตอบอื่น ๆ ที่ควรจะเป็นเพื่อสามารถให้คะแนนได้เชื่อมั่นยิ่งขึ้นและยังเป็นการตรวจสอบดูว่าโจทย์นั้นชัดเจนพอที่จะให้ตอบตามคำเฉลยไว้หรือไม่ อีกทั้งเป็นการคาดคะเนคะแนนและเวลาให้ตอบที่เหมาะสมอีกด้วย
9. ถ้าข้อสอบมีมากข้อ และยากไม่เท่ากัน ควรจัดเรียงลำดับจากง่ายไปหายาก เพื่อยั่วให้ผู้ตอบเกิดกำลังใจในการตอบ การให้ข้อยากไว้ตอนต้นอาจทำให้เด็กตกใจหรือขวัญเสีย ซึ่งจะกระทบกระเทือนคะแนนข้ออื่น ๆ ได้ ดังนั้นควรเขียนคำชี้แจงให้นักเรียนเลือกตอบข้อใดก่อนก็ได้ เพื่อให้นักเรียนเลือกทำข้อง่ายตามความรู้สึกของนักเรียนเอง
10. ถ้าเป็นไปได้ ควรออกข้อสอบให้มีความยากใกล้เคียงกัน

11. ให้ผู้อื่นที่มีความรู้ในสาขาวิชาที่ออกข้อสอบ อ่านวิพากษ์วิจารณ์ เพื่อตรวจสอบข้อบกพร่อง เช่น เขียนคำถามให้ครอบคลุมสิ่งสำคัญที่ต้องการวัดหรือไม่ คำถามยากพอเหมาะหรือไม่

12. หลีกเลี่ยงวิธีการสอบโดยให้เปิดตำราสอบ เพราะการสอบโดยให้เปิดตำราดูได้นี้ เด็กที่เรียนเก่งอยู่แล้วมีความชำนาญในการใช้หนังสือ มีทักษะในการอ่าน มีความคล่องแคล่ว มีความพยายาม และมีความอดทนสูงกว่าเด็กที่เรียนอ่อน ดังนั้นเด็กเก่งจึงมีแนวโน้มที่จะได้คะแนนสูงกว่าปกติ ส่วนเด็กอ่อนจะมีแนวโน้มที่จะได้คะแนนต่ำกว่าเดิม เพราะแม้แต่ค้นหาคำตอบจนไม่สามารถจะตอบได้ทันตามกำหนดเวลา

13. ไม่ควรออกข้อสอบให้เลือกตอบเพียงบางข้อ เพราะนักเรียนอาจจะเลือกตอบข้อต่างกัน ซึ่งจะก่อให้เกิดความแตกต่างในตัวข้อสอบเอง คะแนนของแต่ละคน จึงขึ้นอยู่กับข้อที่เลือกทำ

เยาวดี วิบูลย์ศรี (2528 : 124) และพวงรัตน์ ทวีรัตน์ (2530 : 113) ให้ข้อควรพิจารณาในการสร้างแบบทดสอบความเรียงไว้คล้าย ๆ กัน ดังนี้

1. ต้องมีการพัฒนาวัตถุประสงค์ของการเรียนการสอนที่ชัดเจน และพิจารณาว่า วัตถุประสงค์และเนื้อหาวิชานั้น ๆ หรือเนื้อหาตอนนั้น ๆ เหมาะสมกับแบบทดสอบถามความเรียงหรือไม่

2. จุดมุ่งหมายในการวัดควรให้ผู้ตอบแสดงความรู้ขั้นสูงของ Bloom คือ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ การประเมินค่า ซึ่งมีความมุ่งหมายพื้นฐานที่เหมือนกัน คือ ความมุ่งหมายที่จะให้ผู้ตอบแสดงทักษะในการจัดระบบระเบียบความรู้ต่าง ๆ ที่เคยมีอยู่ออกมา

3. พิจารณาว่า ผู้สอบมีพื้นฐานด้านความเรียง และเนื้อหาวิชาเพียงพอที่จะเขียนตอบเป็นความเรียงได้ เช่น ไม่ใช้สอบกับเด็กเล็ก ๆ

4. พิจารณาว่า เวลาที่ใช้ในการสอบเหมาะสมหรือไม่ เพราะการเวลาในการคิดและเขียนตอบนานจะก่อให้เกิดปัญหาในการสอบ

นอกจากนี้ยังได้เสนอแนวทางในการสร้างแบบทดสอบความเรียงให้มีคุณภาพ มีหลักการที่สำคัญคือ ทำอย่างไรจึงจะทำให้ ข้อสอบหรือข้อคำถามสะท้อนความจริงของสิ่งที่วัด เช่น ความรู้ ความสามารถ ความรู้สึก หรือทักษะที่แท้จริงของผู้ตอบ ซึ่งการสร้างข้อสอบเพื่อสะท้อนความจริงนั้น อยู่ทั้งการวางแผนและการดำเนินการสร้างที่เป็นระบบ การสร้างข้อสอบที่เป็นระบบต้องอาศัยผู้สร้างที่มีความรู้ในเนื้อหา ตลอดจนการวางแผนการออกข้อสอบที่ดี เขียนข้อสอบที่ตรงประเด็น ตลอดจนตรวจสอบคุณภาพของข้อสอบแต่ละข้อ และได้สรุปขั้นตอนการสร้างข้อสอบความเรียงได้ ดังนี้

ขั้นที่ 1 ตัดสินใจเลือกข้อมูลต่าง ๆ ที่ต้องการ โดยการพิจารณาด้วยว่าสิ่งที่กำลังวัดนั้นเหมาะสมกับข้อสอบความเรียงหรือไม่ เช่น การระบุจุดมุ่งหมายในการทดสอบ การระบุเนื้อหาให้ชัดเจน การทำตารางแสดงความสอดคล้องเนื้อหากับจุดมุ่งหมายในการทดสอบ การกำหนดน้ำหนักของข้อสอบ จำนวนข้อ

ขั้นที่ 2 พิจารณาความเป็นอิสระในการตอบที่ยอมให้ในการตอบข้อสอบความเรียงนี้ มีการกำหนดขอบเขตการตอบซึ่งอาจทำได้โดยการจำกัดเนื้อหา จำกัดความยาว จำกัดเวลา จำกัดทรัพยากรที่ใช้ เช่น ให้อภิปรายเฉพาะเหตุผลในตำรา หรือจำกัดรูปแบบการตอบ เช่น ให้อภิปรายการ เขียนแผนภูมิ เขียนโครงร่าง

ขั้นที่ 3 กำหนดรูปแบบหรือลักษณะของข้อคำถาม กำหนดเวลาในการสอบ กำหนดคะแนน เขียนข้อคำถาม เสนอคำตอบ เพื่อประโยชน์ในการตัดสินให้คะแนน

ขั้นที่ 4 ตรวจสอบข้อสอบที่เขียนขึ้น ทดลองใช้ แก้ไขปรับปรุง

ฮอปกินส์ และแอนเตส (Hopkins and Artes, 1990 : 233-234) เสนอข้อแนะนำในการสร้างและพัฒนาแบบทดสอบความเรียงไว้ดังนี้

1. ใช้แบบทดสอบความเรียงเมื่อต้องการคำตอบของนักเรียนที่เป็นการอธิบาย การเปรียบเทียบ ความเหมือน การเปรียบเทียบความแตกต่าง การบรรยาย การวิเคราะห์ การสังเคราะห์หรือประเมินผล
2. ไม่ใช่แบบทดสอบแบบความเรียงในการวัดความจำ
3. ใช้ถ้อยคำในการตั้งคำถามที่ชัดเจนและกระชับ
4. ตั้งคำถามด้วยสถานการณ์ใหม่ ๆ ที่นักเรียนไม่คุ้นเคย ไม่เช่นนั้นนักเรียนก็จะจำสิ่งที่เรียนมาแล้วนั้นมาตอบแบบเดิม ๆ
5. กำหนดขอบเขตในการตอบที่ต้องการไว้ในข้อสอบด้วย และขีดจำกัดในการตอบควรกำหนดตามเนื้อหาที่นักเรียนได้เรียน นอกจากนี้ความยากของคำตอบที่ต้องการก็อาจจะกำหนดโดยเวลา จำนวนข้อ หรือปริมาณที่ว่างในกระดาษคำตอบสำหรับเขียนตอบ
6. บอกให้นักเรียนรู้ว่า จะจัดการกับศัพท์เฉพาะหรือศัพท์เทคนิคในเนื้อหาที่จะสอบอย่างไร นักเรียนควรที่จะต้องให้คำจำกัดความหรืออธิบายคำเหล่านี้ในการตอบด้วยหรือไม่
7. เขียนข้อสอบเป็นข้อ ๆ เพื่อวัดด้านต่าง ๆ ที่ได้ศึกษามา และไม่ควรออกข้อสอบที่ให้ผู้สอบนำข้อมูลทั่วไปมาตอบ
8. ถ้อยคำที่ใช้ในข้อสอบหรือคำสั่งจะต้องมีความหมายชัดเจน
9. ให้นักเรียนทุกคนตอบข้อสอบด้วยชุดข้อสอบชุดเดียวกันและเหมือนกัน
10. เลือกใช้แบบทดสอบแบบความเรียงแบบขยายความ เพื่อวัดทักษะการเขียน การสื่อสารหรือการจัดองค์ประกอบเท่านั้น

ลวัน สายยศ และอังคณา สายยศ (2539 : 86-88) ได้เสนอหลักการสร้างข้อคำถามแบบความเรียงไว้ 5 ข้อ ดังนี้

1. จุดประสงค์ของการสอบก่อน แล้วจึงเขียนข้อคำถาม
2. ควรใช้คำถามที่มีความกระชับชัด ด้วยหลักการถามและหลักภาษา ทั้งนี้เพื่อให้ผู้สอบเข้าใจผิดหรือมัวตีความหมายของคำถามที่เป็นไปได้หลายแง่หลายมุม
3. คำถามหนึ่ง ๆ ควรเป็นเรื่องเดียว ทั้งนี้เพื่อให้ผู้ตอบสามารถตอบตรงเป้าหมายที่ผู้ถามต้องการ
4. คำถามควรคำนึงถึงเวลาให้ผู้ตอบทำการตอบ
5. คำถามทุกคำถามผู้สอบควรทำเฉลยไว้ และวางแผนการให้คะแนนแต่ละส่วนว่าเป็นเท่าไรในเวลาที่จำกัดไว้ คำตอบที่สมบูรณ์ที่สุดควรเป็นอย่างไร เพื่อใช้เปรียบเทียบ

เวอร์ต (Ward, 1999 : 130) ให้คำแนะนำสำหรับการเขียนแบบทดสอบความเรียงไว้ 5 รูปแบบ ดังนี้

1. ตั้งคำถามเกี่ยวกับความคิดรวบยอดที่สำคัญ (Ask questions about major concepts.)
2. ให้นักเรียนแสดงความรู้ของเขา , ไม่ใช่รายงานข้อมูลที่มีอยู่แล้ว (Require students to apply their knowledge, not simply to report facts.)
3. ให้นักเรียนแสดงความสามารถของตนอย่างสมบูรณ์ที่สุดเท่าที่จะทำได้ (Define students' task as completely and specifically as possible.)

4. คำถามต้องสั้นและชัดเจนมากกว่าที่จะเป็นคำถามทั่ว ๆ ไป (Ask many short, specific questions, rather than a few general questions.)
5. แจ้งหัวข้อที่จะให้คะแนนให้ผู้สอบทราบก่อนทำข้อสอบทุกครั้ง (Decide in advance which variables will be considered in scoring and let students know.)

2.3 องค์ประกอบที่ส่งผลต่อคะแนนในแบบทดสอบความเรียง

บุญเชิด ภิญโญนันพงษ์ (2521 : 125) ได้อธิบายเกี่ยวกับความลำเอียงในการให้คะแนนของผู้ตรวจให้คะแนนที่อาจจะเกิดขึ้นโดยไม่ตั้งใจ เช่น

1. The Halo Effect เป็นความลำเอียงชนิดหนึ่งที่เป็นผลจากผู้ตรวจให้คะแนน ได้รับอิทธิพลมาจากผู้เขียนตอบ อันเป็นความประทับใจของผู้ตรวจให้คะแนนที่ได้ประสบจากผู้เขียนตอบนั้น จะทำให้ผู้ตรวจให้คะแนนประเมินค่าคะแนนนั้นออกมาสูงขึ้นหรือต่ำลง เนื่องจากความประทับใจในลักษณะเด่น – ด้อย ของผู้เขียนตอบ ตัวอย่างเช่น เพียงพร เป็นนักเรียนที่ประพฤติดี เคารพขนบอนุรักษ์ ขยันทำการบ้าน คุณลักษณะเหล่านี้จะมีอิทธิพลต่อการตรวจให้คะแนนคำตอบเพียงพรสูงขึ้น ส่วนสันติเป็นนักเรียนที่ขยัน ดื้อกระด้าง เป็นที่ระอาใจของครูอยู่เสมอ การตรวจให้คะแนนคำตอบของสันติซึ่งตอบได้ดีเท่าเพียงพร มีแนวโน้มที่จะได้คะแนนต่ำกว่าของเพียงพร

2. อารมณ์ของผู้ตรวจให้คะแนน เป็นต้นเหตุอีกอย่างหนึ่งที่ทำให้คะแนนผิดเพี้ยนไป ตัวอย่างเช่น ครูคนหนึ่งกำลังตรวจข้อสอบนักเรียนอยู่ มีโทรเลขจากกรุงเทพมหานครมาบอกว่าสอบเข้าเรียนประสานมิตรได้ ครูคนนี้จะตื่นเต้นดีใจ คะแนนจะมีแนวโน้มสูงขึ้นกว่าปกติ ในทางตรงข้าม ถ้าครูคนนี้ได้รับหนังสือจากจังหวัดลงโทษลดขั้นเงินเดือน เพราะไม่แต่งเครื่องแบบข้าราชการ ในวันที่รัฐมนตรีศึกษาเดินทางมาตรวจราชการ อารมณ์ของครูคนนี้จะขุ่นมัวทันที คะแนนจะมีแนวโน้มต่ำลงกว่าปกติ

พวงรัตน์ ทวีรัตน์ (2530 : 119) กล่าวว่า การตรวจให้คะแนนของแบบทดสอบความเรียง มีสิ่งที่ทำให้เกิดความคลาดเคลื่อนได้ ซึ่งสรุปส่วนที่สำคัญมี 4 ประการ คือ

ผลจากสิ่งอื่น (The Halo Effect) ได้แก่ การที่ผู้ตรวจให้คะแนนอ่านคำตอบข้อแรกของผู้สอบคนใดคนหนึ่ง ถ้าผู้สอบตอบได้ดีมาก ก็จะทำให้คะแนนสูง และเมื่อตรวจข้อต่อไป ก็มีแนวโน้มจะให้คะแนนสูงตามไปด้วย แม้ว่าคำตอบข้ออื่น ๆ จะไม่ดีเท่าที่ควร และในทางตรงข้าม ถ้าคำตอบข้อแรก ผู้สอบตอบไม่ค่อยดี ก็จะได้คะแนนต่ำ และข้อต่อ ๆ ไป ก็จะมีแนวโน้มได้คะแนนต่ำตามไปด้วยเช่นกัน นอกจากนั้นความรู้จักมักคุ้นก็เป็นผลต่อการให้คะแนน เช่น ถ้าครูทราบว่ากระดาษคำตอบนั้นเป็นของนักเรียนที่เรียนดี และขยัน ก็จะทำให้คะแนนสูงกว่าที่ควรจะเป็น และในทางตรงข้ามถ้าเป็นของนักเรียนที่เรียนอ่อน และเป็นนักเรียนเกเร ก็จะทำให้คะแนนต่ำกว่าเป็นจริง

ความใจดี (Generosity) ของครู ครูหลายคนมักจะให้คะแนนผู้สอบหรือนักเรียนสูงกว่าที่ควรจะเป็นเสมอ เช่น ให้เกรด A เกือบครึ่งห้องทั้ง ๆ ที่ตามความเป็นจริงแล้ว อาจมีคนทำคะแนนได้ถึงเกณฑ์ที่จะให้เกรด A เพียงครึ่งห้อง หรือ เพียงหนึ่งในสามของนักเรียนทั้งห้องเท่านั้น ความคลาดเคลื่อนที่ขึ้นได้เกิดจากองค์ประกอบข้อนี้คล้ายคลึงกับข้อแรก

การปรับโทษ (Penalty) การให้คะแนนกด หรือการตัดทอนคะแนนนักเรียนที่ไม่สามารถตอบได้ตามที่ครูได้เป้าหมายไว้ ครูก็จะทำการปรับโทษ โดยให้คะแนนต่ำกว่าที่ควรจะเป็น ทั้ง ๆ ที่นักเรียนทำคะแนนอยู่ในระดับที่ใช้ได้ ทั้งนี้เพราะครูได้ตั้งเป้าหมายไว้สูงเกินไป เมื่อนักเรียนตอบได้ไม่ถึงเกณฑ์ที่ครูตั้งเป้าหมาย

ก็จะเกิดความรู้สึกว่านักเรียนมีความสามารถต่ำ จึงให้คะแนนน้อย ถ้าเกิดกรณีเช่นนี้ขึ้น ก็ควรจะโต้ขอให้มีการตรวจสอบคำตอบกับเกณฑ์มาตรฐานวิธีสอน ตลอดจนระดับสติปัญญาของนักเรียนอีกครั้ง ก็จะแก้ไขความคลาดเคลื่อนกรณีนี้ได้

องค์ประกอบแทรกซ้อน (Extraneous Factors) ความคลาดเคลื่อนของการให้คะแนน อาจเกิดจากองค์ประกอบแทรกซ้อนอื่น ๆ เช่น อาจเกิดจากตัวนักเรียนเอง เป็นต้นว่า ลายมือ ความถูกต้องของการเขียน ตัวสะกดการันต์ การเรียบเรียงถ้อยคำ สำนวนภาษา หลักไวยากรณ์ สิ่งเหล่านี้มีผลต่อการให้คะแนนเป็นอย่างมาก ซึ่งครูจะต้องชี้แจงให้นักเรียนทราบอย่างชัดเจนว่า การให้คะแนนนั้นคำนึงสิ่งใดบ้างเพื่อให้นักเรียนได้เพิ่มความระมัดระวัง

เวอร์ต (Ward. 1999 : 162-164) ได้แบ่งปัญหาในการให้คะแนนโดยผู้ตรวจออกเป็น 2 ด้าน คือ ด้านความคลาดเคลื่อนทางเนื้อหา และด้านความคลาดเคลื่อนจากผู้ตรวจ ความคลาดเคลื่อนนี้จะต้องทำให้เกิดขึ้นน้อยที่สุด วิธีที่ช่วยได้ก็คือ การเขียนกำหนดเกณฑ์การตรวจให้ชัดเจนที่สุด

ความคลาดเคลื่อนทางเนื้อหา แบ่งเป็น 2 ชนิด คือ

1. ความคลาดเคลื่อนจากความใจดี (Generosity Error) เป็นผลที่ผู้ตรวจจะให้คะแนนสูง ซึ่งจะพบในหลาย ๆ สถานการณ์ เช่น โดยทั่วไปครูมากกว่า 90 % มักจะให้คะแนนลูกศิษย์ของตนเองสูง
2. ความคลาดเคลื่อนจากความลำเอียง (Halo Effect) เป็นพฤติกรรมส่วนบุคคลและตัวแปรที่เป็นความประทับใจส่วนตัวต่อนักเรียน หรือต่อลายมือที่สวยงาม จะเป็นผลให้คะแนนที่ได้สูงกว่า เช่น ซาร่า เป็นนักเรียนที่ดี (หรือที่น่ารัก , เป็นคนใจดี) ก็จะส่งผลให้คะแนนของเธอสูงขึ้นในทุก ๆ หัวข้อ ทั้ง ๆ ที่เธออาจจะไม่มีความสามารถในด้านนั้น ๆ ก็ได้

ความคลาดเคลื่อนจากผู้ตรวจ มี 2 ลักษณะ คือ

1. ความคลาดเคลื่อนภายในผู้ตรวจ (Intrajudge Errors) เป็นความคลาดเคลื่อนที่เกิดจากการใช้มาตรฐานของผู้ตรวจเองที่อาจเกิดจากความเหนื่อยล้า ความเบื่อ หรือความฉุนเฉียว จากตัวกระดาษคำตอบขณะทำการตรวจให้คะแนน ผู้ตรวจต่างคนก็จะมีมาตรฐานต่างกัน
2. ความคลาดเคลื่อนจากภายนอกผู้ตรวจ (Interjudge Errors) เป็นความคลาดเคลื่อนจากความแตกต่างระหว่างผู้ตรวจต่างคนกัน เช่น ผู้ตรวจคนที่ 1 อาจจะให้คะแนนสูงกับนักเรียนทุกคน และอีกคนอาจจะให้คะแนนต่ำกับนักเรียนทุกคน

3. กฎเกณฑ์การให้คะแนน (Scoring Rubric)

3.1 ความหมายของกฎเกณฑ์การให้คะแนน

กรมวิชาการ (2539 : 54-59) ได้ให้ความหมายไว้ว่า แนวทางในการให้คะแนน (Scoring Guide) ซึ่งจะต้องกำหนดเป็นมาตรวัด (Scale) และรายการคุณลักษณะที่บรรยายถึงความสามารถในการแสดงออกของแต่ละจุดในมาตรวัดอย่างชัดเจน จึงมีส่วนสำคัญในการส่งเสริมการเรียนรู้ โดยทำให้เป้าหมายการแสดงออกของนักเรียนมีความชัดเจนขึ้น นำไปสู่การบรรลุจุดประสงค์หรือสมรรถภาพที่สำคัญของมาตรฐานการศึกษาได้

เลาเวนี เกเรียร์ (2540 : 159) กล่าวว่า เป็นเครื่องมือในการให้คะแนนที่ประกอบด้วยประเด็นต่าง ๆ (Criteria) ที่จะใช้พิจารณางานหนึ่ง ๆ และคำอธิบายระดับคุณภาพของแต่ละประเด็นประเมิน ซึ่งอาจเรียงลำดับตั้งแต่ดีเลิศไปจนถึงต้องปรับปรุง หรือให้เป็นระดับตัวเลขตั้งแต่มากที่สุด (เช่น 4) ไปจนถึงน้อยที่สุด (เช่น 0) ประเด็นประเมินอาจกำหนดเพิ่มเติมได้หลายข้อ คำอธิบายระดับคุณภาพควรอธิบายให้ชัดเจนที่สุด กระชับที่สุด เป็นคำอธิบายที่สามารถบอกได้ว่า ทำไมจึงต้องดีเลิศ ดี ต้องปรับปรุง

สเทนมาร์ค (Stenmark. 1991 : 24) ได้อธิบายถึงกฎเกณฑ์การให้คะแนน (Scoring Rubric) ที่ใช้สำหรับคำถามแบบปลายเปิด (Open-ended Questions) ซึ่งปรับปรุงมาจาก California Assessment Program Grade 12 Generalized Rubric ไว้ว่า กฎเกณฑ์การให้คะแนนเฉพาะ (Specific Rubric) เป็นแนวทางในการให้คะแนนเฉพาะของคำถามแบบปลายเปิดในแต่ละข้อ สำหรับกฎเกณฑ์การให้คะแนนทั่วไป (General Rubric) จะเป็นตัวแบบของกฎเกณฑ์การให้คะแนนเพื่อใช้พิจารณากฎเกณฑ์เฉพาะ รายละเอียดจะเปลี่ยนแปลงสำหรับระดับคะแนนที่แตกต่างกัน แนวคิดพื้นฐานสามารถประยุกต์ใช้ได้ในทุกระดับ

คอร์ทนี (Courtney. 1996 : 77-79) อธิบายว่า เป็นคำแนะนำที่เกี่ยวกับการให้คะแนน ถ้าระบบการให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนนที่มี 5 ระดับ นักเรียนจะได้ 5 คะแนนเต็ม ก็ต่อเมื่อสามารถทำได้ครบทุกอย่างที่ระบุไว้ กฎเกณฑ์การให้คะแนน (Scoring Rubric) เน้นการแสดงความสามารถของนักเรียนในด้านความเข้าใจในความคิดรวบยอด ความรู้เกี่ยวกับกระบวนการหรือขั้นตอน ทักษะในการแก้ปัญหา การใช้เทคโนโลยีใหม่ ๆ (เช่น คอมพิวเตอร์) ในการประมวลผลข้อมูล และการสื่อความหมายกับผู้อื่น

บุญเชิด ภิญโญอนันตพงษ์ (2544 : 90) ให้ความหมายของกฎเกณฑ์การให้คะแนนว่า เป็นชุดของแนวทางในการให้คะแนนผลการปฏิบัติเรื่องใดเรื่องหนึ่ง สำหรับใช้ประเมินคุณภาพการปฏิบัติงานของผู้เรียน แนวทางในการให้คะแนนนั้น อาจทำในรูปของมาตราประเมินค่า หรือแบบตรวจสอบรายการ

จากผลการศึกษาดังกล่าว ผู้วิจัยขอให้ความหมายของ กฎเกณฑ์การให้คะแนน (Scoring Rubric) หมายถึง แนวทางที่กำหนดขึ้นเพื่อชี้บอกระดับของพฤติกรรม คุณภาพของกระบวนการทำงานและผลลัพธ์ หรือสิ่งที่ต้องการ ซึ่งได้มาจากการพิจารณาผลของผู้ตรวจให้คะแนนทำให้เกิดความเข้าใจตรงกัน กฎเกณฑ์การให้คะแนนมีส่วนสำคัญในการส่งเสริมการเรียนรู้ ทำให้เป้าหมายการแสดงผลงานของนักเรียนมีความชัดเจนขึ้น นำไปสู่การบรรลุจุดประสงค์หรือสมรรถภาพที่สำคัญของมาตรฐานการศึกษาได้

3.2 ประเภทของกฎเกณฑ์การให้คะแนน

นิตโก (Nitko. 1983 : 266) ได้จัดประเภทของกฎเกณฑ์การให้คะแนน (Scoring Rubric) ออกเป็น 3 ประเภท คือ

1. Holistic Rubrics เป็นประเภทที่ประเมินหรือให้คะแนนด้านผลผลิตหรือกระบวนการทำงานได้ในภาพรวม ไม่มีการจำแนกเป็นองค์ประกอบเฉพาะ เป็นรูปแบบที่สามารถนำไปใช้ได้ง่าย ลดเวลาในการประเมินต่อนักเรียน 1 คน เมื่อผู้ประเมินนำไปใช้ จะรายงานผลการประเมินคุณลักษณะของการปฏิบัติงานในภาพรวม

2. Analytic Rubrics ประเภทนี้ใช้เวลาในการประเมินมากกว่าในแบบ Holistic Rubrics ผู้ประเมินจะต้องพิจารณาและประเมินทีละด้าน รายละเอียดเหล่านี้เป็นสิ่งที่มีความสำคัญ ต่อการนำไปใช้หรือวินิจฉัยเด็กให้เข้าใจถึงสิ่งที่คาดหวังในการปฏิบัติแต่ละด้าน มีประโยชน์ช่วยในการเรียนรู้ของนักเรียนแม้จะใช้เวลามากก็ตาม

3. Annotated Holistic Rubrics เป็นประเภทที่จำกัดขอบเขตของ Holistic Rubrics และ Analytic Rubrics รวมเอาคุณสมบัติของ Holistic Rubrics และ Analytic Rubrics เข้าด้วยกัน วิธีการของ Annotated Holistic Rubrics คือการประเมินการปฏิบัติในภาพรวมแบบ Holistic Rubrics ก่อน แล้วจึงเลือกเพียง 1 หรือ 2 ประเด็น เพื่อมาประเมินเชิงวิเคราะห์ในแบบ Analytic Rubrics ประโยชน์ของการประเมินประเภทนี้ช่วยให้ผู้ประเมินสามารถประเมินได้เร็วขึ้น สนับสนุนการประเมินในจุดเด่น 2-3 ประเด็น ออกมาให้ชัดเจนได้ จุดเด่นเหล่านี้สามารถนำมาใช้เป็นข้อมูลย้อนกลับ แต่จะไม่มีประโยชน์ในการวินิจฉัย เนื่องจากไม่ได้ประเมินทุกด้าน

ชัยฤทธิ์ ศิลาเดช (2540 : 29) กล่าวว่า การให้คะแนนของรูบรีค เป็นการตอบคำถามว่านักเรียนทำอะไรได้สำเร็จ มีระดับความสำเร็จในขั้นต่าง ๆ กัน หรือมีผลงานอย่างไร และได้แบ่งการให้คะแนนแบบรูบรีคไว้ 2 แบบ คือ

1. แบบรวมองค์ประกอบ (Holistic Score) คือ การให้คะแนนงานชิ้นใด ชิ้นหนึ่งโดยดูภาพรวมของชิ้นงานว่า มีความเข้าใจในความคิดรวบยอด การสื่อความหมาย กระบวนการที่ใช้ และผลงานเป็นอย่างไร แล้วเขียนอธิบายคุณภาพของงาน หรือความสำเร็จของงานเป็นขั้น ๆ โดยอาจจะแบ่งระดับคุณภาพ ตั้งแต่ 0-4 หรือ 0-6

2. แบบแยกองค์ประกอบ (Analytic Score) เพื่อให้การมองคุณภาพของงานหรือความสามารถของนักเรียนได้อย่างชัดเจน จึงได้มีการแยกองค์ประกอบของการให้คะแนน และอธิบายคุณภาพของงานในแต่ละองค์ประกอบเป็นระดับโดยทั่วไปแล้วจะมีการแยกองค์ประกอบของงานเป็น 4 ด้าน คือ

2.1 ความเข้าใจในความคิดรวบยอด ข้อเท็จจริง เป็นการแสดงให้เห็นว่านักเรียนเข้าใจในความคิดรวบยอด หลักการในการแก้ปัญหาที่ถามอย่างกระจ่างชัด

2.2 การสื่อความหมาย สื่อสาร คือ ความสามารถในการอธิบาย นำเสนอ การบรรยาย เหตุผล แนวคิดให้ผู้อื่นเข้าใจได้ดี มีความคิดสร้างสรรค์

2.3 การใช้กระบวนการและยุทธวิธี สามารถเลือกยุทธวิธี กระบวนการนำไปสู่ความสำเร็จอย่างมีประสิทธิภาพ

2.4 ความสำเร็จของงาน ความถูกต้องแม่นยำในผลสำเร็จของงาน หรืออธิบายที่มา และการตรวจสอบผลงาน

บุญเชิด ภิญญอนันตพงษ์ (2544 : 90) กล่าวว่า การพิจารณาคุณภาพเกณฑ์การให้คะแนน ถ้าพิจารณาตามลักษณะของงานหรือกิจกรรม แบ่งออกเป็น 2 รูปแบบ คือ กฎเกณฑ์การให้คะแนนแบบทั่วไป (General Rubric) สำหรับใช้ประเมินงานหรือกิจกรรมหลาย ๆ ชิ้นที่อยู่ในกรอบเนื้อเรื่องเดียวกัน และกฎเกณฑ์การให้คะแนนแบบเฉพาะ (Specific Rubric) เป็นแนวทางในการให้คะแนนทั่วไป ที่เขียนให้เฉพาะเจาะจงกับการปฏิบัติงานหรือกิจกรรมแต่ละชิ้น ถ้าพิจารณาตามเกณฑ์หรือองค์ประกอบของการให้คะแนน ก็ยังแบ่งออกได้เป็น 2 ลักษณะ คือ เกณฑ์รวม (Holistic) เป็นการพิจารณาผลงานของผู้เรียนในภาพรวมว่ามีคุณภาพในระดับใด และเกณฑ์ย่อย (Analytic) เป็นการพิจารณาผลงานของผู้เรียนแยกเป็นรายด้าน

สำหรับกฎเกณฑ์การให้คะแนน (Scoring Rubric) ที่ใช้ในการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ผู้วิจัยได้เลือกใช้รูปแบบเฉพาะ (Specific Rubric) กำหนดระดับคุณภาพเป็น 4 ระดับ ในแต่ละทักษะ

3.3 การสร้างกฎเกณฑ์การให้คะแนน

การสร้างกฎเกณฑ์การให้คะแนนมีขั้นตอนในการดำเนินการ 7 ขั้นตอน (โครงการพัฒนาทรัพยากรมนุษย์. ม.ป.ป.) ดังนี้

ขั้นที่ 1 วิเคราะห์ผลการเรียนรู้ของนักเรียน

วิเคราะห์ผลการเรียนรู้ของนักเรียนในแต่ละจุดประสงค์ หรือ แต่ละหัวเรื่อง (Theme) เมื่อนักเรียนได้ปฏิบัติกิจกรรมการเรียนรู้แล้วจะเกิดผลการเรียนรู้อะไรบ้าง ผลการเรียนรู้ของนักเรียนไม่จำเป็นจะต้องมีครบทุกประเภท ทุกครั้งที่นักเรียนได้ปฏิบัติกิจกรรมการเรียนรู้ บางครั้งในการวางแผนการสอน ครูคาดหวังในผลการเรียนรู้ของนักเรียนเพียงกระบวนการอย่างเดียว บางครั้งครูคาดหวังทั้งกระบวนการและผลงาน

ตัวอย่าง

การพูดอภิปรายเกี่ยวกับเรื่องราวต่างๆ ในชีวิตประจำวัน

ผลการเรียนรู้ของนักเรียนที่ครูคาดหวัง

- 1.การวางแผนการจัดอภิปราย
- 2.การกำหนดเค้าโครงและรูปแบบการอภิปราย
- 3.การพูดอภิปรายที่ดี

ขั้นที่ 2 กำหนดประเด็นที่จะต้องประเมิน

การกำหนดประเด็นที่ต้องการประเมิน อาจให้นักเรียนร่วมกันเสนอความคิดเห็นในการกำหนดประเด็นประเมินผลการเรียนรู้แต่ละอย่างของเขาเอง เช่น ครูกำหนดว่า ถ้าเราจะดูคนดีจะดูอย่างไรบ้าง ให้นักเรียนช่วยกันเสนอประเด็นที่จะดูคนดี ซึ่งมีประเด็นมากมายเป็นของนักเรียนเอง แต่อย่างไรก็ตามเนื่องจากในการวางแผนการสอน ครูผู้สอนได้มีการคาดหวังความสำเร็จจากการเรียนรู้ของนักเรียนเอาไว้ล่วงหน้าแล้ว จึงควรกำหนดรายการประเมินที่สำคัญๆ ของความสำเร็จจากการเรียนแต่ละด้านเอาไว้ล่วงหน้าด้วย เพื่อจะได้แจ้งให้นักเรียนทราบและเพื่อการตรวจสอบผลงานของนักเรียนเอง

ตัวอย่าง ประเมินการพูดอภิปราย

รายการประเมิน

- 1.การแสดงความคิดเห็น
- 2.เนื้อหาสาระ
- 3.การนำเสนอประเด็นอภิปราย
- 4.การใช้ถ้อยคำ
- 5.มารยาทในการพูด

6.การรักษาเวลา

ขั้นที่ 3 คัดเลือกประเด็นประเมินผลที่สำคัญ

การกำหนดประเด็นที่จะต้องประเมิน เราจะพบว่า ในผลการเรียนรู้หนึ่งอย่าง จะมีประเด็นที่ต้องการประเมินมาก โดยเฉพาะถ้าเปิดโอกาสให้นักเรียนส่วนร่วมในการเสนอประเด็นประเมินด้วยแล้ว จะมีมุมมองในการประเมินที่หลากหลาย หรือรวมประเด็นที่สำคัญๆ

ขั้นที่ 4 เลือกรูปแบบในการสร้างเกณฑ์

การประเมินการสร้างเกณฑ์การประเมินผลการเรียนรู้ของนักเรียน เพื่อพิจารณาตัดสินให้คะแนน (Scoring Rubric) มีรูปแบบในการสร้าง 2 รูปแบบ

รูปแบบที่ 1 การให้คะแนนแบบรวมองค์ประกอบ (Holistic Scoring Rubric)

การสร้างเกณฑ์การประเมินแบบรวมองค์ประกอบ หมายถึง การให้คะแนนผลการเรียนรู้โดยรวมทุกประเด็นที่กำหนดเพื่อการประเมิน แล้วเขียนอธิบายคุณภาพของผลการเรียนรู้แต่ละระดับ

ตัวอย่าง รูปแบบการให้คะแนนเป็นภาพรวม

ประเมิน ความสามารถในการอ่านจับใจความสำคัญ

ประเด็นประเมิน

1. การตอบคำถามจากเรื่องที่อ่าน
2. การบอกความสำคัญของเรื่องที่อ่าน
3. การข้อคิดจากเรื่องที่อ่าน
4. การเสนอความคิดเห็นจากเรื่องที่อ่าน

ตาราง 4 เกณฑ์การให้คะแนนแบบรวมองค์ประกอบ

ระดับคุณภาพ	คำอธิบาย
ระดับ 3	หมายถึง ตอบคำถามจากเรื่องที่อ่านถูกทุกข้อ บอกเนื้อหาสาระถูกต้องได้ใจความต่อเนื่อง บอกข้อคิดได้ตรงประเด็นสมบูรณ์ เสนอความคิดเห็นด้วยเหตุผลและประโยชน์
ระดับ 2	หมายถึง ตอบคำถามจากเรื่องที่อ่านผิดไม่เกิน 3 ข้อ จาก 5 ข้อ บอกเนื้อหาสาระได้ถูกต้อง แต่วกวน บอกข้อคิดได้ตรงประเด็น แต่ไม่ต่อเนื่อง เสนอความคิดเห็นด้วยเหตุผล
ระดับ 1	หมายถึง ตอบคำถามจากเรื่องที่อ่านผิดมากกว่า 3 ข้อ ใน 5 ข้อ บอกเนื้อหาสาระได้บ้าง บอกข้อคิดได้บ้าง แต่วกวน เสนอความคิดเห็นแต่ไม่แสดงเหตุผล

รูปแบบที่ 2 การให้คะแนนแบบแยกองค์ประกอบ (Analytic Scoring Rubric)

การให้คะแนนแบบแยกองค์ประกอบ หมายถึง การให้คะแนนโดยการแยกองค์ประกอบของสิ่งที่จะประเมิน เพื่อให้มองเห็นคุณภาพของงาน หรือความสามารถของนักเรียนได้อย่างชัดเจน ผลการประเมินจะบ่งบอกถึงจุดเด่น จุดด้อยของแต่ละประเด็นประเมินได้ชัดเจน การสร้างเกณฑ์การประเมินในรูปแบบนี้จะต้องเขียนคำอธิบายคุณภาพของงานในแต่ละองค์ประกอบ และแต่ละระดับขององค์ประกอบให้ชัดเจน

ตัวอย่าง รูปแบบการให้คะแนนแบบแยกองค์ประกอบ

ประเมิน ความสามารถในการอ่านจับใจความสำคัญ

ประเด็นประเมิน

1. การตอบคำถามจากเรื่องที่อ่าน
2. การบอกความสำคัญของเรื่องที่อ่าน
3. การบอกข้อคิดจากเรื่องที่อ่าน
4. การเสนอความคิดเห็นจากเรื่องที่อ่าน

ตาราง 5 เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ

ประเด็นประเมิน	คำอธิบายระดับคุณภาพ		
	3 (ดี)	2 (พอใช้)	1 (ควรปรับปรุง)
1. การตอบคำถามจากเรื่องที่อ่าน	- ตอบคำถามถูกต้องทุกข้อ	- ตอบคำถามผิดไม่เกิน 3 ข้อ จาก 5 ข้อ	- ตอบคำถามผิดมากกว่า 3 ข้อ จาก 5 ข้อ
2. การบอกความสำคัญของเรื่องที่อ่าน	- บอกเนื้อหาสาระ ถูกต้องได้ใจความต่อเนื่อง	- บอกเนื้อหาสาระ ถูกต้อง แต่วกวน	- บอกเนื้อหาสาระได้บ้าง
3. การบอกข้อคิดจากเรื่องที่อ่าน	- บอกข้อคิดได้ตรงประเด็นสมบูรณ์	- บอกข้อคิดได้ตรงประเด็นแต่ไม่ต่อเนื่อง	- บอกข้อคิดได้บ้างแต่วกวน
4. การเสนอความคิดเห็นจากเรื่องที่อ่าน	- เสนอความคิดเห็นด้วยเหตุผลและเป็นประโยชน์	- เสนอความคิดเห็นด้วยเหตุผล แต่มีประโยชน์น้อย	- เสนอความคิดเห็นแต่ไม่แสดงเหตุผล

ขั้นที่ 5 กำหนดค่าระดับคุณภาพในการประเมิน

ค่าระดับคุณภาพ คือ ตัวเลขที่บ่งบอกถึงคะแนนการประเมินผลการเรียนรู้ของนักเรียน ซึ่งผู้สอนเป็นผู้กำหนด หรืออาจจะให้นักเรียนมีส่วนร่วมในการกำหนดค่าระดับคุณภาพก็ได้ ตัวเลขค่าระดับคุณภาพอาจจะมีระดับ 0-1-2-3 หรือ 0-1-2-3-4 หรือ 0-3-4-5-6 หรือ 0-10-15-20 ในกรณีที่กำหนดค่าระดับต่ำสุดที่เลข 0 นั้นหมายถึง นักเรียนไม่มีผลการเรียนรู้ หรือไม่มี ผลงานเข้ารับการประเมินหรือไม่ได้ดำเนินการตามรายการประเมินนั้น และการกำหนดค่าระดับจาก 0-10 แสดงว่าผู้กำหนดค่าระดับคุณภาพ พิจารณาแล้วเห็นว่า คุณภาพของงานหรือสิ่งที่จะประเมินควรมีคะแนนสูงสุดคือ 10 เช่น การกำหนดค่าระดับในการประเมิน

การแข่งขันทักษะทางวิชาชีพ ระดับเขตการศึกษา ผู้เข้าแข่งขันเป็นผู้ที่ได้รับการคัดเลือกจากจังหวัดมาแล้ว ถือว่าเคยผลิตผลงานที่มีคุณภาพมาก่อน

ขั้นที่ 6 บรรยายคุณภาพการประเมินแต่ละระดับ

เมื่อได้เลือกรูปแบบในการกำหนดเกณฑ์การให้คะแนน และกำหนดตัวเลขค่าระดับแล้ว จะต้องเขียนคำอธิบายขอบข่ายการพิจารณาตัดสินให้คะแนนแต่ละระดับคุณภาพให้ชัดเจน ผลงานหรือกระบวนการในลักษณะไหน อย่างไร จึงมีคุณภาพระดับ 4 ลักษณะไหนอย่างไรจึงมีคุณภาพระดับ 3 ระดับ 2 และระดับ 1 คำบรรยายคุณภาพแต่ละระดับ จะเกิดประโยชน์สูงสุด ถ้าครูผู้สอนสามารถจัดทำขึ้นมา และได้มอบให้กับนักเรียนไปล่วงหน้าพร้อมกับมอบภาระงานให้กับนักเรียน เพราะคำตอบบรรยายคุณภาพแต่ละระดับ จะเป็นเครื่องมือให้นักเรียนได้ตรวจสอบและพัฒนาผลงาน หรือกระบวนการ ทำงานของตนเอง หรือของกลุ่มได้ตลอดเวลา ผลงานของนักเรียนจะเป็นผลงานที่ดี ขึ้นอยู่การที่ครูเขียนคำบรรยายคุณภาพแต่ละระดับเอาไว้ล่วงหน้า จะทำให้นักเรียนมีส่วนร่วมในการสร้างเกณฑ์การประเมิน ครูสามารถนำเอาข้อเสนอของนักเรียนมาปรับปรุงคำอธิบายในตอนหลังก็ได้ ซึ่งก็คงไม่แตกต่างจากที่ครูได้กำหนดไว้เดิมมากนัก

ตาราง 6 คำอธิบายคุณภาพ “การพูดอภิปราย”

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง มีความคิดริเริ่มสร้างสรรค์ มีการนำเสนอข้อเท็จจริงและแสดงเหตุผลอย่างเหมาะสม มีการยกตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย การออกเสียงถูกต้องชัดเจน มีมารยาทในการพูด และรักษาเวลาได้
3 ดี	หมายถึง มีความริเริ่มสร้างสรรค์ มีการนำเสนอข้อเท็จจริง และแสดง เหตุผลอย่างเหมาะสม แต่ขาดตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย ส่วนการออกเสียงถูกต้องชัดเจน มีมารยาทในการพูดและรักษาเวลาได้ดี
2 ใช้ได้	หมายถึง ไม่มีความคิดริเริ่มสร้างสรรค์ มีการนำเสนอข้อเท็จจริงและ ข้อคิดเห็น แต่ขาดการยกตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย มีการออกเสียงถูกต้องชัดเจน แต่ไม่รักษาเวลาและมารยาทในการพูด
1 ใช้ไม่ได้	หมายถึง มีการนำเสนอข้อเท็จจริงและข้อคิดเห็น แต่ขาดการยกตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย การออกเสียงไม่ถูกต้องชัดเจน ขาดการรักษาเวลาและมารยาทในการพูด
0	หมายถึง ไม่มีการอภิปราย

ขั้นที่ 7 กำหนดคะแนนการตัดสินระดับคุณภาพ

การประเมินผลงานหรือชิ้นงานเรามักจะได้ยินคำพูดเชิงการบ่งบอกถึงการจัดระดับคุณภาพ (เกรด) ของผลงาน เช่นพูดว่า ผลงานชิ้นนี้ดีมาก ชิ้นนี้ดี ชิ้นนี้พอใช้ได้ไม่เคยบอกว่า คะแนนระหว่างเท่าไรถึงเท่าไรที่แสดงว่าผลงานดีมาก การกำหนดคะแนนการตัดสินระดับคุณภาพ เป็นการกำหนดช่วงคะแนนจากการประเมินผลการเรียนรู้

ตาราง 7 การกำหนดคะแนนการตัดสินระดับคุณภาพ

คะแนน	ระดับคุณภาพ
9 – 12	ดี
5 – 8	พอใช้
1 – 4	ควรปรับปรุง

ชัยฤทธิ์ ศิลาเดช (2540 : 68) ได้เสนอลำดับขั้นตอนการสร้างกฎเกณฑ์การให้คะแนน (Rubric) ไว้ดังนี้

1. กำหนดขั้นตอนหรือลักษณะเด่นของผลงานที่คาดหวังไว้ตามจุดมุ่งหมาย
2. จัดหัวข้อรายการที่มีความสำคัญและแสดงออกถึงการบรรลุจุดมุ่งหมายไว้อย่างชัดเจน
3. คัดเลือกเฉพาะรายการที่สามารถสังเกตได้หรือประเมินตัดสินได้
4. นำรายการที่เลือกไว้มากำหนดเป็นข้อกระทงในการประเมิน โดยขยายให้ชัดเจนมากขึ้นในลักษณะของพฤติกรรมที่สังเกตได้ หรือคุณลักษณะที่เด่นชัดของผลงาน
5. สร้างเกณฑ์การประเมินในลักษณะของมาตราจัดอันดับคุณภาพ ซึ่งประกอบด้วยเกณฑ์การประเมินผลงานในแต่ละทักษะย่อย
6. นำเกณฑ์การประเมินไปให้ผู้เชี่ยวชาญทางรวบรัดผลพิจารณาความเที่ยงตรงเชิงเนื้อหาตามวิธีของโรวินเนลลีและแฮมเบิลตัน
7. เลือกตัวอย่างผลงานของนักเรียนที่ไม่ได้เลือกเป็นผลงานดีเด่นมาทักษะละ 1 ผลงาน นำไปให้ผู้ตรวจให้คะแนนจำนวน 2 คน ทดลองใช้เกณฑ์การประเมินผลงาน ประเมินให้คะแนนผลงานจนครบทุกทักษะ
8. หาความสอดคล้องของการให้คะแนนโดยใช้ค่าสหสัมพันธ์แบบเพียร์สัน และทดสอบความมีนัยสำคัญทางสถิติด้วย t-test

เสาวนีย์ เกรียร์ (2540 : 159) ได้เสนอกระบวนการสร้างกฎเกณฑ์การให้คะแนน (Scoring Rubric) ที่เน้นการมีส่วนร่วมของนักเรียน ดังต่อไปนี้

1. ดูงานแบบต่าง ๆ ให้นักเรียนดูตัวอย่างของงานที่ดี และชิ้นงานที่ไม่ค่อยดี แล้วบอกถึงลักษณะว่าชิ้นที่ดีมีลักษณะอย่างไร ที่ไม่ดีมีลักษณะอย่างไร
2. กำหนดแนวทาง จากประเด็นแรกให้ช่วยกันตรวจสอบดูว่า งานที่ดีนั้นจะต้องดูอะไรบ้าง
3. คำอธิบายระดับคุณภาพต่าง ๆ ในแต่ละแนวทางนั้นให้อธิบายว่า ระดับสูงสุดมีลักษณะอย่างไร แล้วเติมส่วนที่อยู่ตรงกลางลงไป
4. ฝึกหัดกับงานตัวอย่าง ให้นักเรียนลองใช้รูปกับงานที่ให้ดูเป็นตัวอย่างในขั้นที่ 1
5. หัดใช้ในการประเมินตนเองและให้เพื่อนประเมิน ให้นักเรียนทำ ในระหว่างที่นักเรียนกำลังทำงาน ให้หยุดเป็นระยะ ๆ เพื่อประเมินตนเองและให้เพื่อนประเมิน
6. ทบทวนให้อีกโอกาส นักเรียนทบทวนและปรับปรุงงานของเขาโดยดูจากผลสะท้อนที่ได้จากขั้นตอนที่ 5
7. ครูประเมิน ครูใช้รูปที่นักเรียนใช้นั้น เพื่อประเมินผลงานของนักเรียน

บุญเชิด ภิญโญนันตพงษ์ (2544 : 94) กล่าวว่า การสร้างกฎเกณฑ์การให้คะแนนการประเมินผลภาคปฏิบัติต้องมีส่วนประกอบในการดำเนินการที่สำคัญ 3 ขั้นตอน คือ

1. กำหนดการปฏิบัติให้ชัดเจน

- 1.1 ระบุชนิดของการปฏิบัติว่า เป็นทักษะวิธีการทำหรือผลงานสำเร็จที่สร้างขึ้น
- 1.2 กำหนดจุดสนใจของการประเมินว่า สนใจเป็นรายบุคคลหรือรายกลุ่ม
- 1.3 กำหนดเกณฑ์การปฏิบัติที่สะท้อนคุณลักษณะที่สำคัญของการปฏิบัติที่ประสบผล

สำเร็จ

2. ออกแบบกิจกรรมสำหรับใช้ประเมิน ดังนี้

2.1 ระบุลักษณะของงานหรือกิจกรรมว่า เป็นผลงานหรือการปฏิบัติตามที่มอบหมายให้ทำ หรือเป็นการแสดงออกที่เกิดขึ้นตามธรรมชาติ

2.2 กำหนดรายการของงานหรือกิจกรรม ได้แก่ นิยามจุดหมายของการปฏิบัติ เงื่อนไขสถานการณ์ของการปฏิบัติ และเกณฑ์มาตรฐาน

2.3 กำหนดจำนวนชิ้นงานหรือกิจกรรมให้ทำว่าจะใช้เพียงงานเดียวหรือหลายงานจึงจะเพียงพอในสภาพการณ์นั้น

3. ให้คะแนนและจัดบันทึกผล ดังนี้

3.1 ระบุรูปแบบการให้คะแนนที่ใช้เกณฑ์ในการให้คะแนนแบบใด เช่น แบบทั่วไปหรือแบบเฉพาะเนื้อหา และแบบเกณฑ์รวมหรือแบบเกณฑ์ย่อย

3.2 กำหนดวิธีการจัดบันทึกว่าเป็นแบบตรวจสอบรายการ มาตรฐานประเมินค่าการปฏิบัติ การจัดบันทึกพฤติกรรมตามช่วงเวลา หรือการจัดจำรรณา

3.3 ระบุผู้ประเมินว่าจะให้ใครเป็นผู้สังเกตและประเมิน ผลการปฏิบัติจะเป็นประโยชน์แก่ครูผู้สอน ครูที่เคยสอน ตัวผู้เรียนเอง หรือเพื่อน ๆ ในกลุ่ม

การศึกษาค้นคว้าในครั้งนี้ ผู้วิจัยได้ใช้วิธีผสมผสานจากหลายแนวทางที่กล่าวมาข้างต้น มาสร้างเป็นกฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ โดยเลือกใช้รูปแบบเกณฑ์เฉพาะ (Specific Rubric) จำนวน 5 ทักษะ ตามทักษะย่อยของทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

3.4 แนวทางในการตรวจให้คะแนน

ไพศาล หวังพานิช (2528 : 96-97) กล่าวว่า ไม่ว่าจะใช้วิธีใดตรวจให้คะแนน การตรวจที่จะช่วยเพิ่มความเชื่อถือได้ของคะแนน ควรกระทำดังต่อไปนี้

1. ไม่ควรให้เด็กเขียนชื่อในกระดาษคำตอบ เพื่อป้องกันการให้คะแนนจากความรู้สึกประทับใจในเรื่องอื่น ๆ ของเด็ก ซึ่งเรียกว่า halo effect เช่น คะแนนจากความคุ้นเคย ความรู้สึกว่าเด็กมีความตั้งใจและขยันขันแข็ง เป็นต้น

2. ตรวจคำตอบทีละข้อ ไม่ควรตรวจทุกข้อของแต่ละคน เพราะจะก่อให้เกิด halo effect เช่นกัน เช่น เห็นว่าข้อแรก ๆ ของเด็กได้คะแนนมาก ข้อต่อไปจึงให้คะแนนน้อย (ทั้ง ๆ ที่คำตอบดี) หรือในทางตรงกันข้าม การตรวจคำถามข้อเดียวกันของทุก ๆ คนให้เสร็จ จะช่วยในแง่การเปรียบเทียบการตอบของเด็กทั้งกลุ่มได้ด้วย อีกทั้งช่วยให้การตรวจแต่ละข้อนั้น ๆ ของแต่ละคนยึดเกณฑ์ที่เหมือนกัน

3. ไม่ควรย้อนไปดูคะแนนของเด็กจากข้อที่ตรวจแล้วในการตรวจข้อต่อ ๆ ไป ทั้งนี้ เพื่อป้องกันไม่ให้คะแนนในข้ออื่น ๆ มีผลกระทบต่อการให้คะแนนในข้ออื่น ๆ

4. ไม่ควรให้คะแนนโดยยึดความถูกต้องทางภาษาเป็นหลัก ถ้าหากมิได้มุ่งวัดความถูกต้องในการเขียนและการใช้ภาษา ความถูกต้องสละสลวยในการใช้ถ้อยคำสำนวนในการตอบ ไม่ควรมีอิทธิพลต่อการให้คะแนนมากหรือน้อย ผู้ตรวจควรพิจารณาเฉพาะเป้าหมายการตอบในแง่ความสมบูรณ์ครบถ้วนของเนื้อหาหรือความสมเหตุสมผลของความคิดเป็นหลักในการให้คะแนน

5. ถ้าเป็นไปได้ ควรให้คนอื่นตรวจสอบผลการตรวจให้คะแนน ทั้งนี้เพื่อให้คนอื่นได้ประเมินความเหมาะสมในการให้คะแนนของเรา ตามหลักการที่ถูกต้องนั้น คะแนนที่เด็กได้ควรเป็นคะแนนเฉลี่ยที่เกิดจากผู้ตรวจให้คะแนนหลาย ๆ คน ซึ่งคงเป็นเรื่องยากในเชิงปฏิบัติ แต่อย่างไรก็ตามถ้าไม่อาจปฏิบัติได้อย่างน้อย ผู้ตรวจควรจะได้ทบทวนความเหมาะสมในการตรวจให้คะแนนของตนอีกครั้งหนึ่งก่อนที่จะนำคะแนนเหล่านั้นไปใช้

6. ควรเขียนวิจารณ์ ข้อท้วงติง (Comments) เพื่อประโยชน์ในการเรียนรู้ของผู้เรียน หรืออย่างน้อยได้เป็นหลักฐานว่า ทำไมจึงให้คะแนนเท่านี้ สำหรับคำตอบของเด็กคนนี้

7. การตรวจให้คะแนนต้องกระทำอย่างตั้งใจ โดยเฉพาะอย่างยิ่งต้องมีเกณฑ์การให้คะแนน และต้องอ่านคำตอบของเด็กอย่างถี่ถ้วน เมื่อใดที่ผู้ตรวจขาดหลักเกณฑ์การให้คะแนน หรือไม่ได้อ่านคำตอบอย่างตั้งใจ คะแนนที่ให้กับเด็กมักจะออกมาในรูปกลาง ๆ ซึ่งเป็นไปตามหลักของ Central Tendency Error ซึ่งเป็นไปในลักษณะที่ว่า “เมื่อไม่แน่ใจก็ให้คะแนนกลาง ๆ ไว้ก่อน” ซึ่งแน่นอนคะแนนที่ได้จากการสอบวัดนั้นจะมีความเชื่อมั่นต่ำลง

ส. วาสนา ประवालพฤษย์ (2533 : 39 -42) ได้เสนอวิธีการกำหนดระดับคะแนนทั้งแบบรวมองค์ประกอบ (Holistic Score) ที่มีคะแนนเดียวสำหรับงานหรือข้อเขียนนั้นและแบบแยกองค์ประกอบ (Analytic Score) ที่มีคะแนนหลายคะแนนสำหรับงานหรือข้อเขียนนั้น โดยเสนอวิธีการกำหนดเกณฑ์ 5 วิธี เพื่อใช้ในการตรวจให้คะแนน ดังนี้

1. แยกประเด็นพิจารณาออกเป็นประเด็นย่อย แล้วทำเป็นตารางพิจารณาความถูกต้องในแต่ละประเด็น กำหนดระดับคะแนนตามจำนวนที่ปฏิบัติถูกต้องในประเด็นเหล่านั้น
2. กำหนดระดับความสมบูรณ์ตามเส้นแสดงความต่อเนื่องของความสามารถ (Continuous Ability)
3. กำหนดระดับความผิดพลาด พิจารณาความบกพร่องจากคำตอบว่ามีมากน้อยเท่าใด โดยจะหักจากระดับคะแนนสูงสุดลงมาถึงระดับ
4. กำหนดระดับการยอมรับและคำอธิบาย
5. ใช้หลักการจัดกลุ่มแบบอิงกลุ่ม

สเทนมาร์ค (Stenmark. 1991 : 24) ได้เสนอวิธีการกำหนดเกณฑ์การให้คะแนนเพื่อใช้ในการตรวจให้คะแนนแบบประเมินผลรวม (Holistic) โดยการแบ่งข้อสอบออกเป็นกองต่าง ๆ ดังนี้ ข้อสอบที่มีคะแนนระดับสูงสุด ระดับกลาง และระดับต่ำ และสามารถเพิ่มกองในการแบ่งระดับคะแนนได้ ทุกระดับคะแนนจะอธิบายถึงการประเมินคำตอบในการแก้ปัญหาหนึ่ง ๆ โดยเฉพาะ ไม่นำมาใช้ในการจัดหมวดหมู่ของนักเรียน ระดับคะแนนในแต่ละข้อไม่จำเป็นต้องเหมือนกันกับการสร้างเกณฑ์รูปรีคเฉพาะ (Specific Rubric)

และได้เสนอแนวทางการให้คะแนนแบบรวมองค์ประกอบ (Holistic) โดยแบ่งวิธีการให้คะแนนได้หลายวิธีดังนี้

วิธีที่ 1 แบ่งงานตามคุณภาพเป็น 3 กอง คือ

กองที่ 1 ได้แก่ งานที่มีคุณภาพเป็นพิเศษ และเขียนอธิบายลักษณะของงานที่มีคุณภาพเป็นพิเศษ

กองที่ 2 ได้แก่ งานที่ยอมรับได้และเขียนอธิบายลักษณะของงานที่ยอมรับได้

กองที่ 3 ได้แก่ งานที่ยอมรับได้น้อยหรือยอมรับไม่ได้ และเขียนอธิบายลักษณะของงานที่ยอมรับได้น้อย

จากนั้นก็นำงานแต่ละกองมาให้คะแนนเป็น 2 ระดับ คือ

งานกองที่ 1 จะให้ 6 หรือ 5 คะแนน

งานกองที่ 2 จะให้ 4 หรือ 3 คะแนน

งานกองที่ 3 จะให้ 2 หรือ 1 คะแนน

วิธีที่ 2 กำหนดระดับความผิดพลาด คือพิจารณาความบกพร่องจากคำตอบว่ามีมากน้อยเพียงใด โดยจะหักจากระดับคะแนนสูงสุดลงมาที่ระดับ ดังนี้

คะแนน 4 หมายถึง คำตอบถูก แสดงเหตุผลถูก แนวคิดชัดเจน

คะแนน 3 หมายถึง คำตอบถูก เหตุผลถูก แต่มีข้อผิดพลาดเล็กน้อย

คะแนน 2 หมายถึง เหตุผลหรือการคำนวณผิดพลาด แต่มีแนวทางที่จะนำไปสู่คำตอบ

คะแนน 1 หมายถึง แสดงวิธีคิดเล็กน้อยแต่ไม่ได้คำตอบ

คะแนน 0 หมายถึง ไม่ตอบหรือตอบไม่ถูกเลย

วิธีที่ 3 กำหนดระดับการยอมรับและคำอธิบาย เช่น เกณฑ์การให้คะแนนของความสามารถเข้าใจเนื้อหาสาระ เขียนได้ 4 ระดับดังนี้

คะแนน 4 หมายถึง การสาธิตหรือแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์ ครบถ้วน

ถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนด รวมทั้งเสนอแนวคิดใหม่ que แสดงถึงความเข้าใจอย่างลึกซึ้งถึงกฎเกณฑ์หรือลักษณะของข้อมูล

คะแนน 3 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์ ครบถ้วนและถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงหรือสถานการณ์ที่กำหนด

คะแนน 2 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจที่ไม่สมบูรณ์ ครบถ้วน ถูกต้องในหลักการ ความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนดในบางส่วน

คะแนน 1 หมายถึง การแสดงออกถึงความเข้าใจในหลักการ ความคิดรวบยอด

ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนดได้น้อยมาก และเข้าใจไม่ถูกต้องบางส่วน

คะแนน 0 หมายถึง ไม่แสดงความคิดเห็นใด ๆ

ทั้ง 3 วิธีดังกล่าวนี้ พอที่จะสรุปได้ว่า มีส่วนที่เหมือนกันคือ เป็นการตรวจคะแนนที่มองภาพรวมของผลงาน (ในที่นี้ หมายถึง การตอบคำถามลงในกระดาษคำตอบของนักเรียน) แล้วแยกเป็นกอง ๆ ส่วนที่แตกต่างกันของทั้ง 3 วิธี ก็คือ

วิธีที่ 1 แยกโดยใช้คุณภาพของผลงานเป็นหลัก

วิธีที่ 2 แยกโดยใช้ความผิดพลาดของผลงานเป็นหลัก

วิธีที่ 3 แยกโดยใช้ความสามารถในการอธิบายหรือการแสดงออกถึงความเข้าใจเป็นหลัก

คอทนี (Courtney. 1996 : 34) ได้เสนอลำดับขั้นตอนการให้คะแนนแบบประเมินรวม (Holistic) ไว้ดังนี้

จัดแบ่งงานของนักเรียนออกเป็น 3 กอง

กองที่ 1 งานที่มีคุณภาพพิเศษ

กองที่ 2 งานที่ได้รับการยอมรับ

กองที่ 3 งานที่ไม่ได้รับการยอมรับ

แบ่งงานแต่ละกองออกเป็น 2 ระดับ

กองที่ 1 ให้คะแนน 6 หรือ 5 คะแนน

กองที่ 2 ให้คะแนน 4 หรือ 3 คะแนน

กองที่ 3 ให้คะแนน 2 หรือ 1 คะแนน

งานที่แสดงว่าไม่ได้ใช้ความพยายามเลยให้ 0 คะแนน

และเสนอแนะประเด็นการกำหนดเกณฑ์การให้คะแนนแบบรูปรีดทั่ว ๆ ไป (General Rubric) มีดังนี้

ระดับสูงสุด

- ประกอบด้วยคำตอบที่สมบูรณ์ ชัดเจน ประติดปะต่อ ไม่คลุมเครือและอธิบายได้ดี
- รวมถึงแผนภาพที่ไม่ชัดเจนและซับซ้อน
- สื่อสารอย่างมีประสิทธิภาพเพื่อแสดงต่อผู้อ่าน
- แสดงความเข้าใจในแนวคิดและกระบวนการของคำตอบที่เกี่ยวข้อง
- ระบุส่วนประกอบที่สำคัญของปัญหาทั้งหมด
- รวมถึงตัวอย่างประกอบ
- ให้เหตุผลสนับสนุนที่มีน้ำหนัก
- ความต้องการที่นอกเหนือจากนี้ของปัญหา

ระดับที่สอง

- ประกอบด้วยคำตอบที่ดี แต่อาจมีบางส่วนไม่สมบูรณ์
- อธิบายได้ดีน้อยกว่า สมบูรณ์น้อยกว่า
- ไม่แสดงความต้องการของปัญหานอกเหนือไปจากนี้

ระดับที่สาม

- ประกอบด้วยคำตอบที่สมบูรณ์ แต่การอธิบายสับสน
- แสดงเหตุผลแต่ไม่สมบูรณ์
- รวมถึงแผนภาพไม่เหมาะสม หรือไม่ชัดเจน
- แสดงความเข้าใจในแนวคิดเกี่ยวกับคณิตศาสตร์ แต่ไม่ชัดเจนแน่นอน

ระดับที่สี่

- ยกเว้นส่วนที่น่าสนใจ มีความหมาย หรือทั้งหมดของคำถามและคำตอบ
- มีส่วนใหญ่ผิดพลาด
- ใช้กลยุทธ์ที่ไม่เหมาะสม

ชัยฤทธิ์ ศิลาเดช (2540 : 147) กล่าวว่า การตัดสินให้คะแนนนักเรียนเป็นกิจกรรมที่ประเมินความสามารถของนักเรียนจากหลักฐานผลงานมาตีค่าเป็นคะแนน การที่จะทำให้ผลการตัดสินน่าเชื่อถือนั้น มีปัจจัยสำคัญ 2 ประการ คือ

1. กฎเกณฑ์การให้คะแนน ต้องได้รับการสร้างและพัฒนาจนมีคุณภาพ โดยเฉพาะความเที่ยงตรงตามเนื้อหา นั่นคือ ต้องสร้างให้สอดคล้องกับจุดมุ่งหมายของงาน ในทางปฏิบัติควรช่วยกันสร้างและนำไปทดลองใช้ก่อนเพื่อหาข้อบกพร่องและปรับปรุง

2. ผู้ประเมิน เพื่อให้เกิดความน่าเชื่อถือและยุติธรรมต่อนักเรียน ควรใช้ผู้ประเมินมากกว่า 1 คน และมีการฝึกใช้กฎเกณฑ์การให้คะแนนก่อน เพื่อความชัดเจนและประเมินได้ตรงประเด็น ถ้ามีผู้สอนร่วมกันหลายคน ควรให้ทุกคนประเมินผลงานนักเรียนแต่ละคน แล้วนำเอาคะแนนของผู้ประเมินทุกคนมาเฉลี่ย ในกรณีที่มีความจำเป็นต้องประเมินเพียงคนเดียว ควรประเมินซ้ำอีกครั้งหนึ่ง โดยทิ้งช่วงเวลาการประเมินแต่ละครั้งห่างกันพอสมควร นำคะแนนทั้งสองครั้งมาเฉลี่ย

จากการศึกษาของผู้วิจัยพบว่า ถ้าวางแผนการดำเนินงานชัดเจน มีการกำหนดให้ทำกิจกรรมครบถ้วน ผลสำเร็จของงานสอดคล้องกับกฎเกณฑ์การให้คะแนนแล้ว การตัดสินให้คะแนนของผู้ตรวจจะค่อนข้างสอดคล้องกันมาก ค่าความเชื่อมั่นระหว่างผู้ตรวจค่อนข้างสูง และเป็นไปตามสภาพที่แท้จริงของนักเรียน ผู้วิจัยจึงได้กำหนดรูปแบบของการตรวจให้คะแนนเป็น 2 รูปแบบ แล้วศึกษาค่าความเชื่อมั่นของการตรวจให้คะแนนแต่ละรูปแบบโดยใช้หลักการของทฤษฎีการสุปร้องอ้างอิง

4. ทฤษฎีค่าความเชื่อมั่น (Reliability Theory)

4.1 การประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory)

ทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory) อธิบายความแม่นยำในการวัดโดยใช้ค่าความเชื่อมั่น ยึดข้อตกลงที่เป็นคุณสมบัติคู่ขนานหรือความเท่าเทียมเป็นสำคัญ ซึ่งเป็นไปตามสมมติฐานของสเปียร์แมน (Spearman, 1910) ที่ได้กล่าวถึงข้อตกลงเกี่ยวกับแบบทดสอบคู่ขนานเป็นครั้งแรกว่า เป็นแบบทดสอบทั้งหลายที่วัดคุณลักษณะเดียวกัน

คะแนนที่ได้จากการสอบ (X) ประกอบด้วยคะแนนจริง (T) รวมกับคะแนนความคลาดเคลื่อน (E) จากข้อตกลงของสเปียร์แมน สามารถเขียนเป็นสมการได้ดังนี้

$$X = T + E$$

เมื่อ X แทน คะแนนสังเกต หรือคะแนนที่ได้จากการสอบ (Observed Scores)

T แทน คะแนนจริง (True Scores)

E แทน คะแนนความคลาดเคลื่อน (Error Scores)

คะแนนจริง (True Scores) หมายถึง คะแนนที่ผู้เข้าสอบได้รับ จากการทดสอบด้วยเครื่องมือวัดใด ๆ หรือหมายถึงคะแนนเฉลี่ยของผู้เข้าสอบทำได้จากการสอบหลาย ๆ ครั้ง โดยมีข้อตกลงว่า ต้องไม่มีอิทธิพลจากการฝึกฝน ความเมื่อยล้า และการเรียนรู้ในแบบทดสอบซ้ำ คะแนนความคลาดเคลื่อน (Error Scores) หมายถึง คะแนนที่ผู้เข้าสอบได้หรือไม่ได้เนื่องจากการผิดพลาดที่เกิดจากการวัด ซึ่งมี 2 ชนิด คือ ความคลาดเคลื่อนอย่างสุ่ม (Random Error) เป็นความคลาดเคลื่อนที่เกิดขึ้นโดยบังเอิญ เช่น เด็กบางคนป่วยไม่สามารถทำข้อสอบได้เต็มที่ ทำให้ได้คะแนนลดลง ถ้าความคลาดเคลื่อนอย่างสุ่มเกิดขึ้นในการวัดแล้วไม่มีความสัมพันธ์กับคะแนนจริง นั่นคือสหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนความคลาดเคลื่อนมีค่าเป็น 0 ($r_{TE} = 0$) (Mehrens and Lehmann, 1973 : 91) อีกชนิดหนึ่งที่เรียกว่าความคลาดเคลื่อนที่เป็นระบบ

(Systematic Error) เป็นความคลาดเคลื่อนที่เกิดขึ้นกับนักเรียนทุก ๆ คน เช่น แบบทดสอบที่ใช้พิมพ์โจทย์ตกไป 2 ข้อทุกฉบับ ซึ่งไม่มีผลกระทบต่อค่าความเชื่อมั่น เพราะคะแนนที่คลาดเคลื่อนเกิดขึ้นกับตัวนักเรียนแต่ละคนเหมือนกัน (บุญเชิด ภิญโญนนตพงษ์, 2521 : 270-271)

นอกจากนั้นข้อตกลงเพิ่มเติมเกี่ยวกับคะแนนความคลาดเคลื่อนว่า มีค่าเฉลี่ยเป็น 0 ($\Sigma E = 0$) มีความแปรปรวนเท่ากัน เป็นอิสระต่อกัน และเป็นอิสระต่อคะแนนจริง ภายใต้เงื่อนไขนี้ แบบทดสอบแต่ละฉบับมีค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน และค่าสหสัมพันธ์ระหว่างแบบทดสอบคู่ขนานดังกล่าว มีค่าเท่ากัน เท่ากับอัตราส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนสังเกต และมีความเท่ากันกำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนจริงกับคะแนนสังเกต

ดังนั้นความแปรปรวนของคะแนนสังเกต (σ_X^2) จึงเป็นผลรวมของความแปรปรวนของคะแนนจริง (σ_T^2) กับความแปรปรวนของคะแนนความคลาดเคลื่อน (σ_E^2) เขียนเป็นสมการได้ดังนี้

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

สำหรับสัมประสิทธิ์ความเชื่อมั่น ก็คือ อัตราส่วนระหว่างความแปรปรวนของคะแนนจริง (σ_T^2) กับความแปรปรวนของคะแนนสังเกต (σ_X^2) เขียนเป็นสมการได้ดังนี้

$$\rho_{XX'} = \rho^2_{XT} = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$$

ในระยะเวลาที่ใกล้เคียงกันนั้นบราวน์ (Brown) ได้พัฒนาทฤษฎีความเชื่อมั่น โดยเริ่มด้วยการนิยามแบบทดสอบคู่ขนาน ขณะที่สเปียร์แมนเริ่มศึกษาจากคะแนนสังเกต ประกอบด้วยคะแนนจริงรวมกับคะแนนความคลาดเคลื่อน และจากแนวคิดเกี่ยวกับความเชื่อมั่นของบราวน์ ไปสอดคล้องกับทฤษฎีของสเปียร์แมน (ไพรัตน์ วงษ์นาม, 2533 : 18) จากข้อตกลงเกี่ยวกับคุณสมบัติคู่ขนาน นำไปสู่สูตรการพัฒนาค่าความเชื่อมั่นหลายสูตร รวมทั้งสูตรความสอดคล้องภายใน (Internal Consistency) คือ สูตรสเปียร์แมน-บราวน์ที่ประมาณค่าความเชื่อมั่นจากการสอบครั้งเดียว โดยคำนึงจากค่าสหสัมพันธ์ของแบบทดสอบที่แบ่งเป็น 2 ส่วน ซึ่งอาจแบ่งเป็นข้อคู่-ข้อคี่ หรือ ครึ่งแรก-ครึ่งหลัง ค่าสหสัมพันธ์เมื่อปรับด้วยสูตรแล้ว ค่าที่ได้เป็นค่าประมาณของค่าสหสัมพันธ์ระหว่างแบบทดสอบฉบับนั้นกับแบบทดสอบฉบับอื่นที่มีความคู่ขนานกัน

สูตรหาความเชื่อมั่นแบบสอดคล้องภายใน อีกสองสูตร คือ KR-20 และ KR-21 ใช้ค่าสถิติจากการสอบครั้งเดียว เป็นค่าประมาณสหสัมพันธ์ระหว่าง แบบทดสอบคู่ขนานสองฉบับ ถ้าข้อสอบในแบบทดสอบฉบับนั้นวัดองค์ประกอบเดียว สูตร KR-20 เป็นสูตรที่ให้ค่าสหสัมพันธ์เหมาะสม การใช้สูตร KR-20 ต้องปรับเป็นสูตร KR-21 ถ้าความแปรปรวนของข้อสอบแต่ละข้อมีค่าเท่ากัน สูตรทั้งสองนี้ใช้กับการให้คะแนนแบบ 0 หรือ 1 การให้คะแนนระบบนี้ข้อสอบทั้งฉบับจะต้ององค์ประกอบเดียวกันก็ต่อเมื่อ ข้อสอบมีค่าเฉลี่ยเท่ากันและความแปรปรวนเท่ากัน หมายความว่า สูตรทั้งสองยึดข้อตกลงเบื้องต้นเกี่ยวกับคุณสมบัติคู่ขนาน (Cronbach and others, 1963 : 139-140)

โดยสรุปทฤษฎีความเชื่อมั่นแบบมาตรฐานเดิม พัฒนาจากข้อตกลงของความเท่าเทียม หรือความเป็นคู่ขนานของการวัด จากนิยามความเชื่อมั่นนี้จึงถือว่า เป็นค่าสัมประสิทธิ์การวัดที่มีความเท่าเทียมหรือมีความคู่ขนานกัน จากนิยามนี้ การหาค่าความเชื่อมั่นต้องมีการทดสอบเพื่อให้ได้ค่าจากการทดสอบอย่างน้อยสองชุดแต่ละชุดใช้เครื่องมือวัดที่คู่ขนานกัน คือมีเนื้อหาอย่างเดียวกัน ค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน และค่าสหสัมพันธ์เท่ากัน แล้วคำนวณค่าสหสัมพันธ์ระหว่างค่าที่ได้จากการทดสอบทั้งสองชุด ผลที่ได้เป็นค่าประมาณความเชื่อมั่น กรณีที่ต้องประมาณค่าความเชื่อมั่นจากการสอบครั้งเดียว ต้องแบ่งแบบทดสอบ

เป็นส่วนย่อย อาจแบ่งเป็นสองส่วน เช่น ข้อคู่-ข้อคี่ ครั้งแรก-ครั้งหลัง หรือแบ่งย่อยตามจำนวนข้อ โดยยึดข้อตกลงส่วนย่อยของแต่ละส่วน ต้องมีความเท่าเทียมกัน

แต่ถ้าสถานการณ์ของแบบทดสอบแตกต่างกันไป เทคนิคการประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม ไม่สามารถนำมาประยุกต์ใช้ได้เช่น กรณีของการทดสอบที่มีผู้ตรวจให้คะแนนหลายคน ตรวจให้คะแนนข้อสอบหลายข้อ ของผู้เข้าสอบทุกคนและคะแนนผลการตอบแบบทดสอบที่มีค่าคะแนนต่างกัน เป็นต้น เพราะรูปแบบการวัดของทฤษฎีการทดสอบแบบมาตรฐานเดิมถือว่าคะแนนความคลาดเคลื่อนหรือความคลาดเคลื่อนของการวัดเป็นคะแนนรวมความคลาดเคลื่อนทุกแห่งเข้าด้วยกัน และเป็นความคลาดเคลื่อนรวมอันเดียวกัน ที่ไม่สามารถระบุว่าเป็นอะไรได้ จึงเป็นข้อจำกัดทางทฤษฎี ที่ไม่สามารถศึกษารายละเอียดของแหล่งความคลาดเคลื่อนจากการวัดต่าง ๆ ได้

4.2 การประมาณค่าความเชื่อมั่นที่เกี่ยวข้องกับการตรวจให้คะแนน (Reliability of Rating)

ฮอปกินส์ และ แอนเตส (Hopkins and Antes. 1990 : 550) อธิบายถึงความเชื่อมั่นของผู้ตรวจให้คะแนนที่มีตั้งแต่สองคนขึ้นไป จะมีความเห็นสอดคล้องกันเกี่ยวกับจุดเน้น หรือส่วนสำคัญต่าง ๆ ของคำตอบ ที่ผู้ตรวจให้คะแนนตามความเห็นของผู้ตรวจ ความเชื่อมั่นของผู้ตรวจสูง หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์สูงตามไปด้วย ผู้ตรวจให้คะแนนที่ดีควรเป็นผู้ที่สอนวิชานั้นจริง ๆ (Hopkins and Antes. 1990 : 235)

เจมส์ , เดมารี และ โวล์ฟ (James, Damaree and Wolf. 1993) ได้กล่าวว่า ความเชื่อมั่นที่เกี่ยวข้องกับการตรวจให้คะแนนที่ได้ถือเป็นดัชนีบ่งบอกความสอดคล้อง (Index of Agreement) ไม่ใช่ความคงที่ (Consistency) เขาได้พัฒนาสัมประสิทธิ์ความเชื่อมั่นระหว่างผู้ตรวจ (Inter rater reliability Coefficient , r_{wg}) ซึ่งจะประเมินความสอดคล้องระหว่างผู้ตรวจ โดยกำหนดให้ r_{wg} เป็นการลดสัดส่วนความแปรปรวนที่ผิดพลาดสำหรับการให้คะแนนที่ได้มาเพื่อเทียบกับการจัดแบ่งที่แสดงถึงการตอบแบบสุ่ม (ตัวอย่าง เช่น การแบ่งในกรณีของแต่ละคะแนนในเกณฑ์ของอันดับคะแนนมีความเป็นไปได้เท่า ๆ กันในการนำไปใช้) กล่าวคือ

$$r_{wg} = 1 - (S_x^2 / \sigma_E^2)$$

โดยที่ S_x^2 เท่ากับความแปรปรวนของผลการตรวจให้คะแนนและ σ_E^2 เท่ากับความแปรปรวนของประชากรในการจัดแบ่งอันดับคะแนนที่ได้จากการสุ่ม (เช่น การตรวจให้คะแนนที่เท่ากันของคำตอบที่เป็นไปได้แต่ละประเด็น) ค่าของ r_{wg} มีค่าตั้งแต่ 0.0 จนถึง 1.0 เมื่อความแปรปรวนของคะแนนที่ได้มาเป็นแบบสุ่ม (เช่น $S_x^2 = \sigma_E^2$) ค่า $r_{wg} = 0$ ซึ่งชี้ให้เห็นว่า ไม่มีความสอดคล้องกันเลย หรือ หาก ผลการตรวจให้คะแนนเหมือนกัน (เช่น $S_x^2 = 0$), ค่า $r_{wg} = 1$ ซึ่งชี้ให้เห็นว่ามีความสอดคล้องโดยสมบูรณ์ระหว่างผู้ตรวจทั้งหมด

ค่า r_{wg} ที่ได้ควรจะถูกถือเป็นดัชนีของความเห็นพ้อง (Index of Agreement) ไม่ใช่ค่าความสอดคล้อง (Consistency) เจมส์ (James. 1993) อธิบายว่า การให้คะแนนบางประเภทอาจมีระดับความสอดคล้องสูง (เช่น ผู้ตรวจ 2 คน หรือมากกว่านั้นให้คะแนนออกมาคล้ายคลึงกัน เมื่อให้คะแนนตัวอย่างหลาย ๆ ตัวอย่างบนมาตรฐานเดียวกัน) แต่ค่าการให้คะแนนที่แท้จริงที่ให้ผู้ตรวจคนเดียวอาจจะต่างกันมาก ให้พิจารณาถึงกรณีที่ผู้ตรวจคนที่ 1 ให้คะแนน “4” ทุกครั้ง ผู้ตรวจคนที่ 2 ให้คะแนน “2” และทุกครั้งที่ผู้ตรวจคนที่ 1 ให้คะแนน “3” ผู้ตรวจคนที่ 2 ให้คะแนน “1” ค่าสหสัมพันธ์ระหว่างความคงที่ระหว่างผู้ตรวจ 2 คน นั้นจะสูง

มาก แต่ ระดับความสอดคล้องระหว่างผู้ตรวจนี้จะต่ำ เพราะว่าพวกเขาจะไม่ให้ค่าที่เป็นคะแนนซ้ำกันเหมือนเดิม ข้อดีของสัมประสิทธิ์ rwg ก็คือ สัมประสิทธิ์ที่ได้มาจะแปรผันได้ตามค่าสมมติในการให้คะแนน เช่น เกี่ยวกับการจัดอันดับคะแนนจากการตรวจให้คะแนน

บุญเชิด ภิญโญนันตพงษ์ (2521 : 302) ได้กล่าวถึงการประมาณค่าความเชื่อมั่นในการให้คะแนน (Reliability of rating) หรือค่าความเชื่อมั่นของผู้ตรวจให้คะแนน (Reliability of raters) เมื่อผู้ให้คะแนนมีส่วนเกี่ยวข้องกับคะแนนที่ตรวจ เช่น การให้คะแนนการปฏิบัติหรือการให้คะแนนข้อสอบอัตนัย โดยใช้ผู้ตรวจให้คะแนนมากกว่าหนึ่งคนขึ้นไป ลักษณะเช่นนี้มักจะมีผลคลาดเคลื่อนเกิดขึ้นได้เสมอ เช่น ครูวิทยาศาสตร์ 10 คน ตรวจข้อสอบอัตนัย 5 ข้อ จะพบว่าครูแต่ละคนอาจให้คะแนนแตกต่างกันไป เราจึงต้องตรวจวัดดูว่าครูแต่ละคนให้คะแนนสอดคล้องตามกันหรือไม่เพียงใด ดังนั้นการประมาณค่าความเชื่อมั่นในการให้คะแนนของแต่ละคนจึงเป็นการคำนวณหาความสัมพันธ์ระหว่างผู้ให้คะแนนแต่ละคนจะให้คะแนนสอดคล้องตรงกันมากน้อยเพียงใดนั่นเอง ซึ่งจะเป็นสหสัมพันธ์ภายในกลุ่ม (Intracorelation) การตรวจให้คะแนนอาจจำแนกเป็นสองลักษณะตามผู้ให้คะแนนคือ

1. การตรวจให้คะแนนแบบสมมาตร หมายถึง การที่ผู้ตรวจแต่ละคนให้คะแนนข้อสอบของนักเรียนแต่ละคนครบหมด เช่น นักเรียน 3 คน ครู 2 คน นักเรียนแต่ละคนจะได้รับคะแนนจากครูทั้งสองคน

2. การตรวจให้คะแนนแบบไม่สมมาตร หมายถึง ผู้ตรวจบางคนไม่ตรวจให้คะแนนนักเรียนบางคน เช่น ศาสตราจารย์ทางฟิสิกส์ 8 คน ตรวจให้คะแนนความสามารถทางการวิจัยนิสิตปริญญาโท 28 คน ศาสตราจารย์แต่ละคนขอร้องว่าเขาจะให้คะแนนเฉพาะนิสิตที่เลือกวิจัยในหัวข้อที่ตนรู้เท่านั้น ดังนั้นนิสิตบางคนจะได้รับคะแนนจากผู้ตรวจไม่ครบทุกคน

สำหรับสูตรที่ใช้ในการคำนวณนั้นก็มียุหลายวิธี แต่ยังไม่มีความเห็นสอดคล้องตรงกันว่าสูตรใดจะดีที่สุด

สเนคเดคเคอร์ (บุญเชิด ภิญโญนันตพงษ์, 2521 :303 ; อ้างอิงจาก Snecdecor, 1967. *Statistical methods*.) ได้เสนอสูตรสหสัมพันธ์ภายในกลุ่มใช้กับการตรวจให้คะแนนแบบสมมาตร โดยอาศัยการวิเคราะห์ความแปรปรวน (ANOVA) โดยทำการวิเคราะห์ส่วนประกอบ 3 ส่วน คือ ความแปรปรวนระหว่างนักเรียนระหว่างผู้ตรวจและความคลาดเคลื่อน ซึ่งคำนวณความเชื่อมั่นได้ดังนี้

ความเชื่อมั่นของผู้ให้คะแนนแต่ละคน (Reliability of individual ratings) จะใช้สูตร

$$r_{tt} = \frac{MS_{ss} - MS_E}{MS_{ss} + (k-1)MS_E}$$

เมื่อ r_{tt} แทน ค่าความเชื่อมั่น

MS_S แทน ค่าเฉลี่ยกำลังสอง (ความแปรปรวน) ระหว่างนักเรียน

MS_E แทน ค่าเฉลี่ยกำลังสอง (ความแปรปรวน) ของความคลาดเคลื่อน

ความเชื่อมั่นของการให้คะแนนเฉลี่ยจากแต่ละคน (Reliability of Average ratings) ใช้สูตร

$$r_{tt} = \frac{MS_{ss} - MS_E}{MS_S}$$

การวิเคราะห์ความแปรปรวนคำนวณได้จากสูตรในตาราง ANOVA (บุญเชิด ภิญโญนันตพงษ์, 2521 : 304 ; อ้างอิงจาก Ebel and Mehrens, 1967 : 125. *Measuring Educational Achievement*.) ดังนี้

ตาราง 8 ผลการวิเคราะห์ความแปรปรวนจากการตรวจให้คะแนนแบบสมบูรณ์

Source of Variation	df	Sum of Squares	Mean Square
Total	$Kn-1$	$SS_T = \sum X^2 - \frac{(\sum T)^2}{kn}$	
Rater	$k-1$	$SS_R = k \sum R^2 - \frac{(\sum T)^2}{kn}$	$MS_R = \frac{SS_R}{k-1}$
Students	$n-1$	$SS_S = n \sum T^2 - \frac{(\sum T)^2}{kn}$	$MS_S = \frac{SS_S}{n-1}$
Error	$(k-1)(n-1)$	$SS_E = SS_T - SS_R - SS_S$	$MS_E = \frac{SS_E}{(k-1)(n-1)}$

สำหรับการให้คะแนนแบบไม่สมบูรณ์ แหล่งของผู้ตรวจจะไม่กำหนดให้ ในกรณีนี้ส่วนประกอบของความแปรปรวนจึงมีอยู่สองส่วนคือ ความแปรปรวนระหว่างนักเรียนและความคลาดเคลื่อน ความแตกต่างของการให้คะแนนระหว่างผู้ตรวจแต่ละคนจึงกลายเป็นความคลาดเคลื่อนไป

ในกรณีนี้การวิเคราะห์ความแปรปรวนคำนวณได้จากสูตรในตาราง ANOVA ดังนี้

ตาราง 9 ผลการวิเคราะห์ความแปรปรวนจากการตรวจให้คะแนนแบบไม่สมบูรณ์

Source of Variation	df	Sum of Squares	Mean Square
Total	$K-1$	$SS_T = \sum X^2 - \frac{(\sum T)^2}{k}$	
Between groups	$n-1$	$SS_B = \sum \left[\frac{T^2}{k} \right] - \frac{(\sum T)^2}{k}$	$MS_B = \frac{SS_B}{n-1}$
Within groups	$K-n$	$SS_W = SS_T - SS_B$	$MS_W = \frac{SS_W}{k-n}$

เมื่อ

 K

แทน จำนวนคะแนนทั้งหมด

 k

แทน คะแนนของนักเรียนแต่ละคนที่ได้รับการตรวจจากผู้ตรวจบางคน

 n

แทน จำนวนนักเรียน

การคำนวณความเชื่อมั่นซึ่งจะต้องคำนวณจำนวนผู้ตรวจแต่จำนวนผู้ตรวจของนักเรียนแต่ละคนมีไม่เท่ากัน Snecdecer เสนอแนะให้ใช้ค่าเฉลี่ยแบบฮาร์โมนิค โดยใช้สูตร

$$\tilde{k} = \frac{1}{n-1} \left[\sum k - \frac{\sum k^2}{\sum k} \right]$$

เมื่อ k แทน จำนวนผู้ตรวจ
 n แทน จำนวนนักเรียน

กิลฟอร์ด (บุญเชิด ภิญโญอนันตพงษ์. 2521 : 308 ; อ้างอิงจาก Guilford. 1954. *Psychometric Methods*.) เสนอการประมาณความเชื่อมั่นของการให้คะแนน เมื่อมีผู้ให้คะแนนตั้งแต่สองคนขึ้นไป และให้คะแนนแบบจัดลำดับที่ โดยใช้สูตรดังนี้

$$\bar{r} = 1 - \frac{k(4n+2)}{(k-1)(n-1)} + \frac{12\sum T^2}{k(k-1)n(n^2-1)}$$

เมื่อ $\bar{r}$ แทน ความเชื่อมั่นเฉลี่ยของผู้ตรวจแต่ละคน
 k แทน จำนวนผู้ตรวจ
 n แทน จำนวนนักเรียน
 T แทน ผลรวมของอันดับที่ (Rank) ของนักเรียนแต่ละคน

ถ้าต้องการหาความเชื่อมั่นของผู้ตรวจทั้งหมดปรับขยายความเชื่อมั่นเฉลี่ย โดยใช้สูตรปรับขยายของ Spearman-Brown

4.3 การประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง (Generalizability Theory)

วิธีการหนึ่งที่เป็นประโยชน์ต่อการศึกษาเชิงจำแนกแหล่งความคลาดเคลื่อนของการวัด ได้แก่ การวิเคราะห์ความแปรปรวน การนำเทคนิคการวิเคราะห์ความแปรปรวนซึ่งเป็นวิธีการทางสถิติมาประมาณค่าความเชื่อมั่นของแบบทดสอบและความคลาดเคลื่อนจากการวัดมีมานานแล้ว ในการศึกษาด้านการวัดผลการศึกษา วิธีที่รู้จักกันอย่างแพร่หลาย ได้แก่ วิธีการวิเคราะห์ของฮอยท์ (Hoyt. 1941) ต่อมา ครอนบัคและคนอื่น ๆ (Cronbach and others. 1963) เสนอทฤษฎีการสรุปอ้างอิง (Generalizability Theory) ทางการทดสอบขึ้นเป็นครั้งแรก เพื่อให้ใช้ประมาณค่าความเชื่อมั่นทั่ว ๆ ไปของแบบทดสอบในสถานการณ์ของการทดสอบแบบต่าง ๆ หลังจากนั้น ได้มีนักการศึกษาหลายคนศึกษาทฤษฎีนี้และปรับปรุงวิธีการให้รัดกุมและง่ายขึ้น ผู้ที่มีส่วนพัฒนาทฤษฎีนี้อย่างมากคือ เบรนแนน (Brennan. 1983) ได้ปรับปรุงวิธีการให้ง่ายขึ้นในแง่การตีความและการคำนวณ คาร์ดินีเนทและอัลลาล (Cardinet and Allal. 1983) กล่าวถึงการประยุกต์ทฤษฎีการสรุปอ้างอิงว่า การประยุกต์ใช้ในทางจิตวิทยานั้น มักใช้กับการศึกษาถึงความแตกต่างระหว่างบุคคล สิ่งที่ถูกวัดมักได้แก่ บุคคล องค์ประกอบจึงจำกัดอยู่เฉพาะองค์ประกอบที่สอดคล้องกับเงื่อนไขการวัดเท่านั้น โดยหลักการของทฤษฎีนี้สามารถทำได้กว้างขวางยิ่งขึ้น ดังนั้น คาร์ดินีเนทและคนอื่น ๆ (Cardinet and Allal. 1983) นำหลักการของความสมมาตร (Principle of Symmetry) มาใช้วิเคราะห์การสรุปอ้างอิง หลักการนี้ให้การยอมรับองค์ประกอบแต่ละตัวของรูปแบบการวัดใด ๆ สามารถเลือกให้เป็นสิ่งที่ถูกวัดได้ทั้งสิ้น แล้วแต่จุดมุ่งหมายของการวัดนั้น

4.3.1 โครงสร้างของทฤษฎีการสรุปอ้างอิง

ความเข้าใจเกี่ยวกับทฤษฎีการสรุปอ้างอิงทางการทดสอบเบื้องต้นนั้น จำเป็นต้องทำความเข้าใจ และมีความรู้ในเรื่องพื้นฐานโครงสร้างของทฤษฎีนี้ ในเรื่องต่าง ๆ ดังนี้ (Brennan. 1983 : 2-6)

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 องค์ประกอบที่บ่งชี้แหล่งความแปรปรวนของรูปแบบการวัด ในแบบจำลองการวิเคราะห์ความแปรปรวน รูปแบบความสัมพันธ์นี้ จัดแบ่งออกเป็น 3 ลักษณะ คือ

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

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

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

4. เอกภพ (Universe) หมายถึง เงื่อนไขของการวัดทั้งหมดของแต่ละองค์ประกอบ เช่น จำนวนข้อสอบทุกข้อ จำนวนผู้ตรวจทั้งหมด เอกภพมีความหมายคล้ายกับคำว่า ประชากร (Population) แต่คำว่า เอกภพมักใช้กับองค์ประกอบที่เกี่ยวกับการวัด ส่วนคำว่าประชากรมักใช้กับองค์ประกอบที่เป็นสิ่งที่ถูกวัด

คะแนนเอกภพ (Universe Scores) หมายถึง คะแนนเงื่อนไขการวัดทั้งหมดของแต่ละองค์ประกอบ ซึ่งคล้ายกับคะแนนจริง (Ideal Score) ของสิ่งที่ถูกวัดในทฤษฎีการวัดแบบมาตรฐานเดิม

5. เอกภพของการสังเกตที่ยอมรับได้ (Universe of Admissible Observation) เป็น กลุ่มเงื่อนไขของการวัดที่สามารถวัดหรือสังเกตได้ในแต่ละองค์ประกอบ เช่น ถ้าต้องการวัดผลการเรียนของนักเรียนโดยครูต้องการข้อสอบที่สามารถนำไปใช้ได้ และต้องการผู้ตรวจที่มีคุณภาพ ครูจึงได้กำหนดเงื่อนไขของการวัดเป็นองค์ประกอบของข้อสอบและองค์ประกอบของผู้ตรวจให้คะแนน เรียกว่า เอกภพของการสังเกตที่ยอมรับได้ของครูคนนี้ประกอบด้วย องค์ประกอบของข้อสอบและองค์ประกอบของผู้ตรวจให้คะแนน

6. เอกภพของการสรุปอ้างอิง (Universe of Generalization) หมายถึง เงื่อนไขของการวัดทั้งหมดขององค์ประกอบที่ผู้วิจัยต้องการสรุปอ้างอิงผลการวัด จากกลุ่มตัวอย่างของเงื่อนไขการวัดเหล่านี้ไปยังกลุ่มเงื่อนไขการวัดทั้งหมดขององค์ประกอบนั้น ๆ ดังนั้น เงื่อนไขขององค์ประกอบในเอกภพของการสรุปอ้างอิง

จึงเป็นสับเซตหรือเซตย่อย ๆ ในเอกภพของการสังเกตที่ยอมรับได้ เอกภพของการสุ่มอย่างนี้ ผู้วิจัยต้องระบุประเด็นต่าง ๆ ให้ชัดเจน ซึ่งนำไปใช้ในการออกแบบการวัดในการวิเคราะห์เพื่อการตัดสินใจ (D Study)

7. การศึกษาเพื่อการสุ่มอย่างอิง (Generalizability Study : G Study) หมายถึง การวิเคราะห์การสุ่มอย่างอิง เพื่อประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ ของรูปแบบการวัดโดยเริ่มจากการกำหนดองค์ประกอบต่าง ๆ กำหนดความสัมพันธ์ขององค์ประกอบ และกำหนดเอกภพของการสังเกตที่ยอมรับได้ และมีการวิเคราะห์ความแปรปรวน การประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ

8. การศึกษาเพื่อการตัดสินใจ (Decision Study : D Study) หมายถึง การศึกษาเพื่อประเมินวิธีการวัดที่เหมาะสมเพื่อการตัดสินใจหรือลงข้อสรุป เป็นขั้นสุดท้ายของการวิเคราะห์การสุ่มอย่างอิง โดยเริ่มจากผู้วิจัยต้องระบุสิ่งที่ต้องการวัด เอกภพของการอ้างอิง จำนวนเงื่อนไขของการวัด หรือขนาดของกลุ่มตัวอย่างของแต่ละองค์ประกอบและรูปแบบของการวัด แล้วประมาณค่าความแปรปรวนจากการศึกษา ค่าสัมประสิทธิ์การสุ่มอย่างอิง คือ ความเชื่อมั่นของแบบทดสอบ และสุดท้ายปรับปรุงรูปแบบการวัดเพื่อหาวิธีที่เหมาะสมเพื่อให้เกิดประโยชน์สูงสุดต่อไป

ขนาดของกลุ่มตัวอย่างในการศึกษาเพื่อการตัดสินใจ ไม่จำเป็นต้องมีขนาดเหมือนกับขนาดของกลุ่มตัวอย่าง ที่ใช้ในการศึกษาเพื่อการสุ่มอย่างอิง แต่ขึ้นอยู่กับนำไปใช้เพื่อการตัดสินใจเกี่ยวกับสิ่งที่จะวัดจริง ๆ ซึ่งจะถูกกำหนดขึ้นโดยมีขนาดของกลุ่มตัวอย่างในการศึกษาเพื่อการสุ่มอย่างอิงเป็นพื้นฐาน

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

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

11. สัมประสิทธิ์การสุ่มอย่างอิง (Generalizability Coefficient : E_p^2) หมายถึง อัตราส่วนของความแปรปรวนของคะแนนเอกภพและความแปรปรวนของคะแนนสังเกตที่คาดหวัง ซึ่งสัมประสิทธิ์การสุ่มอย่างอิงที่ประมาณค่าได้จากกำลังสองของค่าสหสัมพันธ์ระหว่างคะแนนเอกภพและคะแนนสังเกต

11.1 สัมประสิทธิ์การสุ่มอย่างอิง สำหรับการวัดที่มีการตัดสินใจในเชิงเปรียบเทียบหรือการวัดแบบอิงกลุ่ม ความคลาดเคลื่อนในการวัด เป็นความคลาดเคลื่อนแบบสัมพัทธ์ ดังนั้น ค่าสัมประสิทธิ์การสุ่มอย่างอิงประมาณได้จาก อัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและผลบวกของความแปรปรวนคะแนนเอกภพกับความแปรปรวนของความคลาดเคลื่อนสัมพัทธ์

11.2 สัมประสิทธิ์การสุ่มอย่างอิง สำหรับการวัดที่มีการตัดสินใจแบบสัมบูรณ์ หรือการวัดแบบอิงเกณฑ์ ความคลาดเคลื่อนในการวัดเป็นความคลาดเคลื่อนแบบสัมบูรณ์ ค่าสัมประสิทธิ์การสุ่มอย่างอิงประมาณค่าได้จาก อัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและผลบวกของความแปรปรวนของคะแนนเอกภพ กับความแปรปรวนของความคลาดเคลื่อนแบบสัมบูรณ์

4.3.2 หลักพื้นฐานของทฤษฎีการสรุปอ้างอิง

1. ข้อตกลงเบื้องต้นของทฤษฎีการสรุปอ้างอิง อาศัยข้อตกลงของแบบการวิเคราะห์ความแปรปรวนที่สัมพันธ์กับรูปแบบของการวัดที่ต้องการศึกษา โดยทั่วไปมีข้อตกลงดังนี้

1.1 ต้องระบุเอกภพที่ต้องการสรุปอ้างอิงให้ชัดเจน จนสามารถบอกได้ว่ามีเงื่อนไขใดบ้างที่เป็นสมาชิกของเอกภพนั้น

1.2 เงื่อนไขการวัดเป็นอิสระต่อกัน กล่าวคือ คะแนนของนักเรียนที่ทำข้อสอบข้อ i ถูกหรือผิดไม่ขึ้นอยู่กับการตอบข้ออื่น

1.3 คะแนนสังเกต (X) สมมติให้เป็นค่าการวัดในมาตราช่วง (Interval Scale)

2. ความหมายของค่าสัมประสิทธิ์การสรุปอ้างอิง

2.1 ค่าสัมประสิทธิ์ของการสรุปอ้างอิง เป็นดัชนีที่สามารถอธิบายความแม่นยำของการวัด เช่นเดียวกับค่าสัมประสิทธิ์ความเชื่อมั่นแบบมาตรฐานเดิมสามารถใช้คำนวณความเชื่อมั่นของคะแนนเอกภพหรือใช้ในการสมการถดถอย ในการประมาณค่าคะแนนเอกภพและใช้ในการปรับแก้ค่าสหสัมพันธ์ที่ลดลงอันเนื่องมาจากความคลาดเคลื่อน (Correcting Correlation for Attenuation) (Cronbach and others. 1963 : 156)

2.2 ค่าสัมประสิทธิ์การสรุปอ้างอิง เป็นค่าประมาณของค่าเฉลี่ยของค่าสหสัมพันธ์ระหว่างค่าการวัดที่สุ่มมาจากเอกภพรายคู่ (Cronbach. 1972 : 157)

2.3 ค่าสัมประสิทธิ์การสรุปอ้างอิง เป็นค่ากำลังสองของสหสัมพันธ์ระหว่างคะแนนเอกภพกับคะแนนสังเกต (Brennan and Kane. 1979 : 40)

2.4 ค่าสัมประสิทธิ์การสรุปอ้างอิง อธิบายได้ในรูปอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพกับคะแนนสังเกต (Cronbach and others. 1972)

3. ข้อแตกต่างระหว่างค่าความเชื่อมั่นกับค่าสัมประสิทธิ์การสรุปอ้างอิง

ถึงแม้ว่าค่าสัมประสิทธิ์การสรุปอ้างอิง มีความหมายเช่นเดียวกับ ค่าความเชื่อมั่นของทฤษฎีการวัดแบบมาตรฐานเดิม แต่มีประเด็นที่แตกต่างกันดังนี้

3.1 การวัดแต่ละครั้งมีค่าสัมประสิทธิ์การสรุปอ้างอิงได้มากกว่า 1 ค่า

3.2 การสรุปอ้างอิงไปยังเอกภพใด ต้องระบุและอธิบายเอกภพนั้นให้ชัดเจนและต้องสุ่มเงื่อนไขนั้นมาศึกษาด้วย

3.3 ค่าสัมประสิทธิ์การสรุปอ้างอิง สามารถบอกถึงความเป็นเอกพันธ์ของเอกภพได้ ถ้าข้อสอบที่นำมาศึกษา เป็นตัวอย่างสุ่มจากเอกภพข้อสอบที่มีความเป็นเอกพันธ์ ก็สามารถใช้คะแนนสังเกตแทนคะแนนเอกภพได้อย่างมั่นใจ

4. การวิเคราะห์สัมประสิทธิ์การสรุปอ้างอิง

แนวคิดในการประมาณค่าพารามิเตอร์ ของทฤษฎีการสรุปอ้างอิง ครอนบัค และคนอื่น ๆ (Cronbach and others. 1972) เบรนนอน (Brennan. 1983) เสนอวิธีการประมาณค่าพารามิเตอร์ เป็น 2 ขั้นตอน ดังนี้

ขั้นที่ 1 การศึกษาเพื่อการสรุปอ้างอิง (G Study) เป็นขั้นแรกของการวิเคราะห์ข้อมูลเพื่อที่จะประมาณค่าความแปรปรวนจากแหล่งต่าง ๆ ซึ่งมีกระบวนการดังนี้

1.1 กำหนดสิ่งที่จะวัด (Object of Measurement) เช่น นักเรียนมัธยมศึกษาปีที่ 4 เป็นต้น

1.2 กำหนดองค์ประกอบ (Facets) ของการวัด เช่น องค์ประกอบข้อสอบ องค์ประกอบผู้ตรวจ เป็นต้น

1.3 กำหนดเอกภาพของการสังเกตที่ยอมรับได้ เช่น การวิจัยครั้งนี้คือ เอกภาพของการสังเกตที่ยอมรับได้ คือ องค์ประกอบข้อสอบ และองค์ประกอบของผู้ตรวจให้คะแนน

1.4 กำหนดความสัมพันธ์ระหว่างองค์ประกอบ (Facets) เช่น ถ้าต้องการศึกษาแบบให้ผู้สอบทุกคนทำข้อสอบทุกข้อ ก็จะเป็นแบบไขว้ (Crossed) คือ $p \times i$ (p คือ นักเรียน และ i คือ ข้อสอบ) แต่ถ้าให้นักเรียนทุกคนทำข้อสอบบางข้อจะเป็นแบบ (Nested) คือ $p : i$

1.5 กำหนดชนิดขององค์ประกอบ (Facets) ว่าเป็นชนิดใดดังต่อไปนี้

1.5.1 องค์ประกอบสุ่มแท้จริง (Purely Random Facets) ได้แก่ องค์ประกอบ (Facets) ที่ทุกเงื่อนไขการวัดได้มาจากการสุ่มอย่างง่ายจากเอกภาพที่มีขนาดไม่จำกัด (∞)

1.5.2 องค์ประกอบสุ่มจำกัด (Finite Random Facets) ได้แก่ องค์ประกอบ (Facets) ที่ทุกเงื่อนไขการวัดได้มาจากการสุ่มอย่างง่ายจากเอกภาพที่มีขนาดจำกัด

1.5.3 องค์ประกอบกำหนด (Fixed Facets) ได้แก่ องค์ประกอบ (Facets) ที่ประกอบด้วยเงื่อนไขการวัดจำนวนจำกัดในเอกภาพและเงื่อนไขการวัดเหล่านี้ถูกเลือกมาศึกษาทั้งหมด

1.6 กำหนดรูปแบบ (Model) การวัดว่าเป็นรูปแบบใดต่อไปนี้

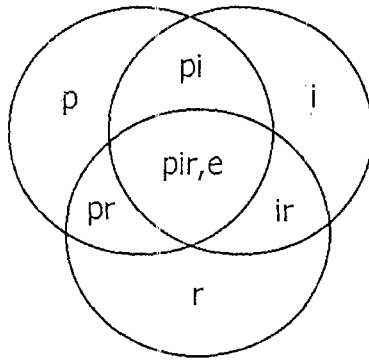
1.6.1 รูปแบบกำหนด (Fixed Model) เป็นรูปแบบที่ประกอบด้วยองค์ประกอบ (Facets) กำหนดทั้งหมด

1.6.2 รูปแบบสุ่ม (Random Model) เป็นรูปแบบที่ประกอบด้วยองค์ประกอบ (Facets) สุ่มแท้จริงและหรือองค์ประกอบ (Facets) สุ่มจำกัด

1.6.3 รูปแบบผสม (Mixed Model) เป็นรูปแบบที่ประกอบด้วยองค์ประกอบสุ่มแท้จริงและ / หรือสุ่มจำกัด และองค์ประกอบ (Facets) กำหนดผสมกัน

การกำหนดสิ่งที่จะวัด องค์ประกอบ เอกภาพของการสังเกตที่ยอมรับได้และความสัมพันธ์ระหว่างองค์ประกอบว่าเป็นแบบ "Crossed" หรือ "Nested" แล้วก็ใช้ ANOVA ทำการวิเคราะห์ให้สอดคล้องกับแบบแผนของความสัมพันธ์ขององค์ประกอบเพื่อประมาณค่า Mean Square แล้วประมาณค่าความแปรปรวนของแหล่งต่าง ๆ เราเรียกค่าประมาณความแปรปรวนของสิ่งที่ถูกวัดว่า ค่าประมาณความแปรปรวนของคะแนนเอกภาพ (Estimated Universe Score Variance) ซึ่งสอดคล้องกับค่าประมาณความแปรปรวนของคะแนนจริง (True Score Variance) ของทฤษฎีการวัดมาตรฐานเดิม เรียกค่าประมาณความแปรปรวนขององค์ประกอบอื่นว่า ค่าประมาณความแปรปรวนของความคลาดเคลื่อน (Estimated Error Score Variance) การวิเคราะห์ชั้น G Study ใช้ค่าขนาดของกลุ่มตัวอย่างที่ศึกษาจริง ๆ แต่ในการวิเคราะห์ชั้น D Study สามารถกำหนดขนาดของกลุ่มตัวอย่างแตกต่างไปจากที่วิเคราะห์ใน G Study ได้

1.7 การเก็บรวบรวมและการวิเคราะห์ข้อมูล เช่น ในรูปแบบสุ่ม (Random Model) โดยการสุ่มกลุ่มตัวอย่างมาศึกษา คือ ผู้สอบ n_p คน ข้อสอบ n_i ข้อ และผู้ตรวจ n_r คน ซึ่งเป็นการศึกษา G Study แบบ $p \times i \times r$ คือ ผู้สอบทุกคนทำข้อสอบทุกข้อและผู้ตรวจแต่ละคนจะตรวจข้อสอบทุกข้อ แล้วคำนวณค่าความแปรปรวนขององค์ประกอบต่าง ๆ วิธีการดังกล่าวนี้จะได้แหล่งความแปรปรวน 7 แหล่ง (Brennan, 1983 : 27 – 28) เขียนแทนด้วยแผนภูมิ Venn ได้ดังภาพประกอบ 1 ดังนี้



ภาพประกอบ 1 แผนภูมิ Venn แสดงองค์ประกอบความแปรปรวนของแบบแผน $p \times i \times r$ เมื่อองค์ประกอบที่ศึกษาทั้งหมดเป็นองค์ประกอบสุ่ม

จากภาพประกอบจะพบว่ามีอิทธิพลหลัก (Main Effects) 3 แหล่งได้จากอิทธิพลจากผู้สอบ (p) อิทธิพลจากข้อสอบ (i) ส่วนผลปฏิสัมพันธ์ (Interaction Effects) 4 แหล่ง คือ ปฏิสัมพันธ์ระหว่างผู้สอบและข้อสอบ (pi) ปฏิสัมพันธ์ระหว่างผู้สอบและผู้ตรวจ (pr) ปฏิสัมพันธ์ระหว่างข้อสอบและผู้ตรวจ (ir) และ ปฏิสัมพันธ์ระหว่างผู้สอบ ข้อสอบ และผู้ตรวจ (pir)

ค่าความแปรปรวนของแหล่งต่าง ๆ เหล่านี้คำนวณจาก Mean Square โดยใช้ ANOVA แบบ Factorial Design $p \times i \times r$ ในการประมาณค่าความแปรปรวนจาก Mean Square นั้นใช้สูตรเฉพาะอย่าง ที่สอดคล้องกับรูปแบบ เช่น สูตรในการประมาณค่าความแปรปรวนของแหล่งต่าง ๆ เมื่อองค์ประกอบที่ศึกษาเป็นองค์ประกอบสุ่มทั้งหมด และกำหนดรูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็นแบบ Crossed ใช้สูตรของ (Brennan. 1983 : 130) ดังนี้

$$\begin{aligned}
 EMS_p &= \sigma^2(pir) + n_i \sigma^2(pr) + n_i \sigma^2(pi) + n_i n_r \sigma^2(p) \\
 EMS_i &= \sigma^2(pir) + n_p \sigma^2(ir) + n_r \sigma^2(pi) + n_p n_r \sigma^2(i) \\
 EMS_r &= \sigma^2(pir) + n_r \sigma^2(pr) + n_p \sigma^2(ir) + n_p n_i \sigma^2(r) \\
 EMS_{pi} &= \sigma^2(pir) + n_r \sigma^2(pi) \\
 EMS_{pr} &= \sigma^2(pir) + n_i \sigma^2(pr) \\
 EMS_{ir} &= \sigma^2(pir) + n_p \sigma^2(ir) \\
 EMS_{(pir)} &= \sigma^2(pir)
 \end{aligned}$$

โครงสร้างของสมการนี้สัมพันธ์กับแผนภูมิ Venn ข้างต้น กล่าวคือ ในวงกลม p ประกอบด้วย พจน์ (Term) ต่าง ๆ 4 พจน์ มี p , pi , pr และ pir เป็นเทอมทางขวามือของสูตรในการหาค่า EMS_p หรือในพื้นที่ pi ซึ่งเป็นส่วนร่วมระหว่างวงกลม p และ i ประกอบด้วย 2 พจน์ คือ pi และ pir ก็คือ

พจน์ที่ปรากฏอยู่ทางขวามือของสูตรในการหาค่า EMS_{pi} นั้นเอง ดังนั้น จึงสามารถสร้างสูตรหาค่า EMS โดยอาศัยแผนภูมิ Venn โดยใช้ n_p เป็นตัวคูณควบ ถ้าองค์ประกอบนั้นไม่รวม p และคูณ n_i ถ้าองค์ประกอบนั้นไม่มี i รวมอยู่ด้วยและ n_r ถ้าองค์ประกอบนั้นไม่มี r รวมอยู่ด้วย

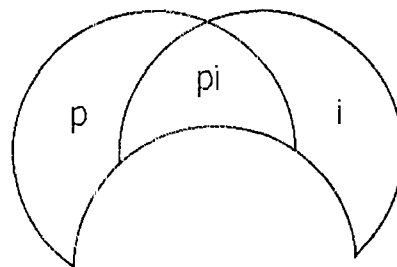
การคำนวณเมื่อองค์ประกอบที่ศึกษาเป็นแบบผสม คือ มีทั้งองค์ประกอบกำหนดและองค์ประกอบสุ่ม และรูปแบบความสัมพันธ์ระหว่างองค์ประกอบเป็นแบบ Crossed สมมติให้ผู้ตรวจ (r) เป็นองค์ประกอบกำหนด สูตรในการคำนวณ รวมทั้งแหล่งความแปรปรวนที่ต้องการประมาณค่าเป็นดังนี้ (Brennan. 1983 : 71 – 72)

$$EMS_p = n_r \sigma^2(pi) + n_i n_r \sigma^2(p)$$

$$EMS_i = n_r \sigma^2(pi) + n_p n_r \sigma^2(i)$$

$$EMS_{pi} = n_r \sigma^2(pi)$$

เนื่องจากองค์ประกอบของผู้ตรวจ (r) เป็นองค์ประกอบที่กำหนด จึงหายไปจากแผนภูมิเดิมเหลือเพียงวงกลม p และ i เท่านั้น ดังแผนภูมิ Venn ในภาพประกอบ 2 ดังนี้



ภาพประกอบ 2 แผนภูมิ Venn แสดงองค์ประกอบความแปรปรวนของแบบแผน $p \times i \times r$ เมื่อ p และ i เป็นองค์ประกอบสุ่ม และ r เป็นองค์ประกอบกำหนด

ค่าความแปรปรวนที่ได้จาก G Study เป็นค่าประมาณความแปรปรวนที่แท้จริงของเอกภพของเงื่อนไขการวัดที่มีขนาด 1 หน่วย กล่าวคือ เป็นความแปรปรวนของข้อสอบ 1 ข้อ ผู้ตรวจ 1 คน ของนักเรียน 1 คน เช่น $\sigma^2(p)$ เป็นค่าประมาณความแปรปรวนของ $\sigma^2(p)$ หมายความว่า จากประชากรที่ทำการวัดนั้น จะได้คะแนนของนักเรียน (p) แต่ละคนจากการสอบข้อสอบทุก ๆ ข้อ (N_p) และได้รับการตรวจจากผู้ตรวจทุกคน (N_r) ในเอกภพของการสังเกตคะแนนเฉลี่ยของนักเรียนแต่ละคนคือคะแนนเอกภพ (μ_p) ดังนั้นความแปรปรวนของ (μ_p) จากการสอบวัดทุก ๆ คนในประชากรคือ $\sigma^2(\mu_p)$ หรืออาจเขียนให้ง่ายว่า $\sigma^2(p)$ เป็นความแปรปรวนของคะแนนเอกภพ ตรงกับความแปรปรวนของคะแนนจริงตามทฤษฎีการวัดมาตรฐานเดิม

ขั้นที่ 2 การศึกษาเพื่อการตัดสินใจสรุปอ้างอิง (D Study) เป็นขั้นการนำผลที่ได้จากการประมาณความแปรปรวนในการศึกษาเพื่อการสรุปอ้างอิง (G Study) มาเลือกวิธีวัดที่เหมาะสมที่สุดภายใต้การตัดสินใจมีขั้นตอนดังนี้

2.1 กำหนดเอกภพของการสรุปอ้างอิง (Universe of Generalization) เป็นเป้าหมายสำคัญของ D Study ได้แก่การกำหนดลักษณะเฉพาะของเอกภพของการอ้างอิงที่ผู้ตัดสินใจต้องการอ้างอิงถึง อาจจะประกอบด้วยเงื่อนไขทั้งหมดในเอกภพของการสังเกตที่ยอมรับได้หรืออาจเป็นเซตย่อย (Subset) ของเอกภพของการสังเกตที่ยอมรับได้ (Brennan. 1983 : 3)

2.2 กำหนดรูปแบบ (Model) หมายถึง จะใช้รูปแบบใด เช่น รูปแบบสุ่ม (Random Model) หรือรูปแบบกำหนด (Fixed Model) หรือรูปแบบผสม (Mixed Model) ซึ่งโดยหลักการแล้วควรใช้โมเดลแบบสุ่มดีที่สุด เพราะจะทำให้การสรุปอ้างอิงได้กว้างขวางกว่าโมเดลแบบผสมและโมเดลแบบกำหนดตามลำดับ

2.3 ขนาดของกลุ่มตัวอย่างของ D Study (D Study Sample Sizes) หมายถึง จำนวนเงื่อนไขขององค์ประกอบใน D Study ไม่จำเป็นต้องเหมือนกัน และสามารถกำหนดได้แตกต่างจากใน G Study โดยใช้สัญลักษณ์ (') แทนขนาดของกลุ่มตัวอย่างใน D Study เช่น n'_i และ n'_r แทน จำนวนข้อสอบ และจำนวนผู้ตรวจ

2.4 กำหนดโครงสร้างของแบบที่ศึกษา (D Study design Structure) โครงสร้างของแบบที่ศึกษาหรือความสัมพันธ์ขององค์ประกอบที่ศึกษา อาจจะเหมือนกันหรือแตกต่างกันจาก G Study ก็ได้

2.5 การประมาณค่าความแปรปรวนในขั้น D Study (Estimate D Study Variance Components) ในขั้นนี้จะมีการประมาณค่าความแปรปรวนขึ้นมาใหม่ ขึ้นอยู่กับการตัดสินใจของนักวิจัยโดยอาศัยผลการประมาณค่าในขั้น G Study เป็นฐานและให้สอดคล้องกับแบบแผนและขนาดกลุ่มตัวอย่างที่ต้องการตัดสินใจ เช่นในแบบแผน $p \times i \times r$ ขนาดตัวอย่างข้อสอบ = n'_i และขนาดของผู้ตรวจ = n'_r ดังแสดงในตารางต่อไปนี้ (Brennan. 1983 : 58)

ตาราง 10 การประมาณค่าความแปรปรวนขึ้นมาใหม่ของแบบแผน $p \times i \times r$

ความแปรปรวนใหม่	ความแปรปรวนเดิม/ขนาดของกลุ่มตัวอย่าง
$\sigma^2(P)$	$\sigma^2(P)$
$\sigma^2(I)$	$\sigma^2(i) / n'_i$
$\sigma^2(R)$	$\sigma^2(r) / n'_r$
$\sigma^2(PI)$	$\sigma^2(pi) / n'_i$
$\sigma^2(PR)$	$\sigma^2(pr) / n'_r$
$\sigma^2(IR)$	$\sigma^2(ir) / n'_i n'_r$
$\sigma^2(PIR)$	$\sigma^2(pir) / n'_i n'_r$

2.6 การประมาณค่าความแปรปรวนของความคลาดเคลื่อน นอกจาก การประมาณค่าองค์ประกอบความแปรปรวน (Variance Components) ตามรูปแบบและขนาดของกลุ่มตัวอย่างที่ต้องการแล้ว ยังจะต้องประมาณค่าความแปรปรวนของความคลาดเคลื่อนซึ่งแบ่งเป็น 2 ชนิด คือ

2.6.1 ความแปรปรวนคลาดเคลื่อนแบบสัมบูรณ์ (Absolute Error Variance)

ใช้สัญลักษณ์ว่า $\sigma^2(\Delta)$ เป็นความแปรปรวนของความแตกต่างระหว่างคะแนนที่ได้จากการสังเกตกับคะแนนที่ได้จากเอกภพของผู้เข้าสอบ คำนวณจากผลบวกของค่าองค์ประกอบความแปรปรวนทั้งหมดยกเว้นความแปรปรวนของคะแนนเอกภพ จากแบบแผน $p \times I \times R$ จะได้

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

2.6.2 ความแปรปรวนคลาดเคลื่อนแบบสัมพัทธ์ (Relative Error Variance)

ใช้สัญลักษณ์ว่า $\sigma^2(\delta)$ หมายถึง ค่าความแปรปรวนของผลต่างระหว่างส่วนเบี่ยงเบนคะแนนที่สังเกตได้และส่วนเบี่ยงเบนคะแนนเอกภพ คำนวณจากผลบวกของค่าองค์ประกอบความแปรปรวนระหว่างสิ่งที่ถูกวัดกับองค์ประกอบอื่น ๆ จากแบบแผน $p \times I \times R$ จะได้สูตรว่า

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

2.7 การประเมินค่าสัมประสิทธิ์การสรุปอ้างอิง เป็นขั้นสุดท้ายของ D Study คือ การคำนวณค่าสัมประสิทธิ์การสรุปอ้างอิง (Generalizability Coefficient) ซึ่งเป็นดัชนีชี้บ่งถึงความเชื่อถือได้ของการวัด (Dependability of Measurement) โดยใช้สูตรคำนวณได้จากอัตราส่วนระหว่างความแปรปรวนของคะแนนเอกภพและความแปรปรวนของคะแนนที่สังเกตที่คาดหวัง

ในการวัดแบบอิงกลุ่ม จะได้สูตรที่ใช้ในการคำนวณค่าสัมประสิทธิ์การสรุปอ้างอิงดังนี้ (Brennan, 1983 : 17)

$$Ep^2 = \frac{\sigma^2(p)}{\sigma^2(p) + \sigma^2(\Delta)}$$

หรือ

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

เมื่อ Ep^2 คือ ค่าสัมประสิทธิ์การสรุปอ้างอิงซึ่งมีความหมายเช่นเดียวกับความเชื่อมั่น

5. งานวิจัยเกี่ยวข้อง

5.1 งานวิจัยเกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์ในต่างประเทศ

แพพพิลิสและคณะ (Pappelis and Others, 1980 : 30-36) ศึกษาการสอนเพื่อปรับปรุงทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนเตรียมแพทย์และเตรียมทันตแพทย์ พบว่าการที่ให้นักศึกษาปฏิบัติกิจกรรมที่ฝึกทักษะกระบวนการทางวิทยาศาสตร์ โดยจัดให้เป็นประสบการณ์ที่นักศึกษาได้พบ สามารถช่วยให้ นักศึกษามีทักษะกระบวนการทางวิทยาศาสตร์สูงขึ้น

พีทัส และฮาเลย์ (Petus and Haley. 1980 : 273-276) ศึกษาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ กับทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน ของนักเรียนที่เข้าค่ายฝึกเยาวชนในรัฐเวอร์จิเนีย สหรัฐอเมริกา จำนวน 505 คน ผลการศึกษาแสดงว่า ตัวแปรต่าง ๆ ซึ่งได้แก่ เพศ อายุ ระดับชั้น ความสนใจที่เกี่ยวกับอาชีพที่ใช้วิทยาศาสตร์และจำนวนหลักสูตรที่นักเรียนเรียนจบ มีความสัมพันธ์กับผลสัมฤทธิ์ด้านการใช้ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน ตัวแปรเหล่านี้สามารถอธิบายความแปรปรวนผลสัมฤทธิ์การใช้ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานได้ร้อยละ 24

กอสไบ (Gosbi. 1982 : 1914 – A) ได้ศึกษาความสัมพันธ์ระหว่างผลสัมฤทธิ์ในการใช้ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานกับพัฒนาการทางสติปัญญาและตัวแปรอื่น ๆ ได้แก่ เกรดเฉลี่ย และความถนัดทางการเรียนวิทยาศาสตร์ กลุ่มตัวอย่างประชากรคือ นักศึกษาที่เลือกเรียนวิชาแคลคูลัสและวิชาวิทยาศาสตร์เป็นวิชาหลัก จำนวน 85 คน พบว่า พัฒนาการทางด้านสติปัญญา เกรดเฉลี่ยและความถนัดทางการเรียนของนักศึกษาทั้งหมด มีความสัมพันธ์กับทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน อย่างมีนัยสำคัญทางสถิติ

พาดิลลา และคณะ (Padilla and Others. 1983 : 239 – 246) ได้ศึกษาความสัมพันธ์ระหว่างทักษะกระบวนการทางวิทยาศาสตร์ผสม และความสามารถในการคิดแบบนามธรรมตามทฤษฎีของเปียเจต์ กลุ่มตัวอย่างเป็นนักเรียนเกรด 7 – 12 จากโรงเรียนนอกเมืองออตลันตาและเจอร์เจีย พบว่าความสามารถในการใช้ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานสัมพันธ์กับการคิดอย่างมีเหตุผล อย่างมีนัยสำคัญทางสถิติ

ฮาร์คูส และเพนิค (Harkoos and Penick. 1983 : 629 – 637) ได้ศึกษาอิทธิพลของบรรยากาศในชั้นเรียนที่มีต่อทักษะกระบวนการทางวิทยาศาสตร์และผลสัมฤทธิ์ทางการเรียน กลุ่มตัวอย่างเป็นนักศึกษาวิทยาลัยดูเพจ (Du Page) ในรัฐอิลลินอยส์ พบว่า บรรยากาศในชั้นเรียนมีอิทธิพลต่อทักษะกระบวนการทางวิทยาศาสตร์ คือ กลุ่มทดลองมีทักษะกระบวนการทางวิทยาศาสตร์สูงกว่ากลุ่มควบคุม แต่ผลสัมฤทธิ์ทางการเรียนของทั้งสองกลุ่มไม่แตกต่างกัน

สตรอวิทซ์ และมาโลน (Strawitz and Malone. 1987 : 56 – 60) ได้ศึกษาเปรียบเทียบความรู้ที่ได้รับและความคงทนของความรู้ในทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานที่ได้รับ จากการเรียนรู้โดยครูสอนและการเรียนรู้ด้วยตัวเอง ผลการศึกษาพบว่า กลุ่มที่เรียนด้วยตัวเองและลงมือปฏิบัติ มีผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์สูงกว่ากลุ่มที่เรียนด้วยวิธีเดิม ที่ไม่มีการลงมือปฏิบัติ และพบว่ามีความคงทนของความรู้ในทักษะกระบวนการทางวิทยาศาสตร์ ขั้นผสม

โดตี (Doty. 1986 : 3311 – A) ได้ทำการศึกษาเปรียบเทียบผลการสอนแบบสืบสวนและการสอนแบบเก่าที่มีผลสัมฤทธิ์ทางการเรียน เจตคติต่อวิทยาศาสตร์และทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานของนักเรียนระดับ 9 และศึกษาความสัมพันธ์ระหว่างเพศ เชื้อชาติ ระดับของสติปัญญาและผลสัมฤทธิ์ทางการเรียน และประสบการณ์ทางวิทยาศาสตร์ที่ผ่านมา กลุ่มตัวอย่างเป็นนักเรียนที่เรียนวิทยาศาสตร์กายภาพในโรงเรียนในเขตมิสซิสซิปปี ผลการศึกษาพบว่า ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานและเจตคติทางวิทยาศาสตร์ ทั้งสองกลุ่มไม่แตกต่างกัน

รูบิน (Rubin. 1989 : 3469 – A) ได้ศึกษายุทธการสอนทำต้นแบบเป็นระบบเพื่อพัฒนาทักษะกระบวนการทางวิทยาศาสตร์แบบบูรณาการ และความสามารถในการใช้เหตุผลในเชิงพุทธพิสัยแบบเป็นทางการ กลุ่มตัวอย่างเป็นนักเรียนระดับกลางในเมือง นักเรียนกลุ่มที่ 1 ได้รับการสอนทักษะกระบวนการทางวิทยาศาสตร์ในช่วงเวลา 3 เดือน จากครูซึ่งได้รับการฝึกกลยุทธ์การสอนการทำต้นแบบเชิงระบบทดลองกลุ่มที่ 2 ได้รับการสอนจากครูที่ฝึกกลยุทธ์การควบคุม วงจรการเรียนรู้ และทักษะกระบวนการทางวิทยาศาสตร์ กลุ่มที่ 3 เป็นกลุ่มควบคุมได้รับการสอนวิทยาศาสตร์ตามแบบเดิม ผลการวิจัยพบว่านักเรียนในกลุ่มที่ 1 และ

กลุ่มที่ 2 มี ผลสัมฤทธิ์ทางการเรียนทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกัน อย่างมีนัยสำคัญทางสถิติ นักเรียนกลุ่มที่ 2 และนักเรียนกลุ่มที่ 3 มีผลสัมฤทธิ์ทางการเรียนทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกัน อย่างมีนัยสำคัญทางสถิติ โดยสรุปกลยุทธ์การสอนการดำเนินงานเป็นระบบ จัดเป็นวิธีการที่มีประสิทธิภาพในการพัฒนาความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์แบบบูรณาการของนักเรียน

ชาร์แมนน์ (Scharmann. 1989 : 715 – 726) ได้ศึกษาเกี่ยวกับอิทธิพลของพัฒนาการของการสอนทักษะกระบวนการทางวิทยาศาสตร์ จากการศึกษาพบว่า การเรียนทักษะกระบวนการทางวิทยาศาสตร์ใน 1 ภาคเรียน จะทำให้พัฒนาการในการรับรู้พื้นฐานทางด้านเนื้อหาวิทยาศาสตร์สูงขึ้น

รอต และรอยเชาต์ฮิวรี (Roth and Roychoudhury. 1993 : 127 – 152) ได้ศึกษาการพัฒนาการทักษะกระบวนการทางวิทยาศาสตร์ในสิ่งแวดล้อมจริง กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษา แบ่งเป็นกลุ่มทดลอง 3 กลุ่ม กลุ่มแรกเป็นนักเรียนเกรด 11 ได้รับการสอนโดยวิธีชี้แนะแนวทางในวิชาฟิสิกส์ กลุ่มที่ 2 เป็นนักเรียนเกรด 12 ได้รับการสอนในวิชาฟิสิกส์ระดับสูง และกลุ่มที่ 3 เป็นนักเรียนเกรด 8 ได้รับการสอนในวิชาวิทยาศาสตร์ทั่วไป โดยที่นักเรียนจากกลุ่มทดลองทั้งสามกลุ่มจะเป็นศูนย์กลางในการเรียน มีการฝึกให้ปฏิบัติการทดลองในห้องปฏิบัติการทดลองอย่างอิสระเหมือนกัน พบว่า กลุ่มทดลองทั้งสามกลุ่มมีพัฒนาการด้านทักษะกระบวนการทางวิทยาศาสตร์สูงขึ้น เช่น มีทักษะในตัวแปรต่าง ๆ เหมือนกัน คือ ทักษะการแปลความหมาย ข้อมูล ทักษะการตั้งสมมุติฐาน ทักษะการให้คำนิยามและทักษะการทดลอง

5.2 งานวิจัยเกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์ในประเทศ

น้อยทิพย์ ศัตรศาสตร์ (2521 : บทคัดย่อ) ได้ศึกษาความสัมพันธ์ระหว่างทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานกับความสามารถในการแก้ปัญหาและผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 4 จำนวน 300 คน นักเรียนชาย 153 คน นักเรียนหญิง 147 คน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล คือ แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน แบบทดสอบการแก้ปัญหา และแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ พบว่า ทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานมีความสัมพันธ์กับความสามารถในการแก้ปัญหาและผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 และคะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานสามารถพยากรณ์คะแนนความสามารถในการแก้ปัญหาและคะแนนผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์

แสงศิริ ศิริมงคล (2529 : บทคัดย่อ) ได้ศึกษาผลสัมฤทธิ์ทางการเรียนและทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่สอนโดยการทบทวนความรู้พื้นฐานด้วยบทเรียนทบทวนและด้วยครูทบทวน พบว่า นักเรียนที่ได้รับการสอนโดยการทบทวนความรู้พื้นฐานด้วยบทเรียนทบทวนมีทักษะกระบวนการทางวิทยาศาสตร์สูงกว่า นักเรียนที่ได้รับการสอนโดยการทบทวนความรู้พื้นฐานด้วยครูอย่างมีนัยสำคัญทางสถิติ

กันยารัตน์ ฤทธิ์บำรุง (2531 : 83) ได้ศึกษาผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนที่เรียนโดยการฝึกอภิปรายแก้ปัญหาร่วมกันทั้งห้อง เป็นกลุ่มย่อยและเป็นรายบุคคล พบว่า ผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนทั้ง 3 กลุ่ม แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ปรีชา เจตนิย (2531 : 63-64) ได้เปรียบเทียบทักษะกระบวนการทางวิทยาศาสตร์และเจตคติเชิงวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีรูปแบบการจัดชั้นเรียนและแผนการเรียนแตกต่างกันใน

โรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา เขตการศึกษา 11 ปีการศึกษา 2530 ในส่วนที่เกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์ พบว่า

1. นักเรียนที่มีแผนการเรียนแตกต่างกัน มีทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

2. นักเรียนที่มีรูปแบบการจัดชั้นเรียนต่างกัน มีทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกันอย่างมีนัยสำคัญทางสถิติ

กาญจนา ลาภรวาย (2532 : 75) ได้ศึกษาผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ด้านทักษะกระบวนการทางวิทยาศาสตร์ จำนวน 7 ทักษะ คือ การสังเกต การวัด การจำแนกประเภท การจัดกระทำและสื่อความหมายข้อมูล การพยากรณ์ การตั้งสมมุติฐาน การตีความหมายข้อมูลและลงข้อสรุปของนักเรียนชั้นประถมศึกษาปีที่ 5 ที่มีระดับความสามารถทางการเรียนวิชาวิทยาศาสตร์ต่างกัน โดยการสอนสาธิตแบบไม่ชี้แนวทางเป็นกลุ่มทดลองและการสอนสาธิตแบบชี้แนวทาง พบว่า

1. นักเรียนกลุ่มทดลองมีผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์สูงกว่านักเรียนกลุ่มควบคุม อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. นักเรียนที่มีระดับความสามารถทางการเรียนสูง ปานกลาง และต่ำ มีผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ญานี ทองทับ (2536 : 134-135) ได้ศึกษาทักษะกระบวนการทางวิทยาศาสตร์ของ นักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดสำนักงานการศึกษากรุงเทพมหานคร ผลปรากฏว่า นักเรียนในโรงเรียนขนาดใหญ่ ขนาดกลาง ขนาดเล็ก มีทักษะกระบวนการทางวิทยาศาสตร์อยู่ในระดับต่ำกว่าเกณฑ์ขั้นต่ำ และนักเรียนที่อยู่ในโรงเรียนเขตเมืองและเขตชานเมือง ก็มีทักษะกระบวนการทางวิทยาศาสตร์อยู่ในระดับต่ำกว่าเกณฑ์ทั้งสองเขต

สุมาลี ตำรังไชย (2537 : 111-121) ศึกษาเปรียบเทียบผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ และทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ที่สอนโดยใช้แบบฝึกทักษะการทดลองจากวัสดุในท้องถิ่นกับการสอนตามคู่มือครู กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 1 ภาคเรียนที่ 2 ปีการศึกษา 2536 ของโรงเรียนคณทิพย์พิทยาคม อำเภอเมือง จังหวัดกำแพงเพชร จำนวน 60 คน แบ่งเป็นกลุ่มทดลองและกลุ่มควบคุม กลุ่มละ 30 คน ผลการศึกษาพบว่า ผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์และทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนกลุ่มทดลองสูงกว่ากลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

มณีรัตน์ เกตุไสว (2540 : บทคัดย่อ) ศึกษาผลการจัดกิจกรรมการทดลองที่มีต่อผลสัมฤทธิ์ทางการเรียนวิชาฟิสิกส์ด้านโมเมนต์ทางวิทยาศาสตร์ และด้านทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ภาคเรียนที่ 1 ปีการศึกษา 2539 โรงเรียนปทุมคงคา เขตคลองเตย กรุงเทพมหานคร จำนวน 67 คน พบว่า

1. ผลสัมฤทธิ์ทางการเรียนวิชาฟิสิกส์ด้านโมเมนต์ทางวิทยาศาสตร์ของนักเรียนกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ผลสัมฤทธิ์ทางการเรียนวิชาฟิสิกส์ด้านโมเมนต์ทางวิทยาศาสตร์เกี่ยวกับทฤษฎีของนักเรียนกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ผลสัมฤทธิ์ทางการเรียนวิชาฟิสิกส์ด้านโมเมนต์ทางวิทยาศาสตร์ที่เกี่ยวกับกลุ่มสัมพันธ์ของนักเรียนกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

4. ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการของนักเรียนกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

วิชชุดา งามอักษร (2541 : บทคัดย่อ) ศึกษาผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ และความสามารถในการคิดอย่างมีเหตุผลของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 โดยการสอนแบบ เอส เอส ซี เอส กับการสอนตามคู่มือครู กลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2540 โรงเรียนบริหารแจ่มใสวิทยา 1 อำเภอดอนเจดีย์ จังหวัดสุพรรณบุรี พบว่า

1. ผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ของนักเรียนกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนกลุ่มควบคุมและกลุ่มทดลองแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ความสามารถในการคิดอย่างมีเหตุผลของนักเรียนกลุ่มควบคุมและกลุ่มทดลอง แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01

มนวิภา อ่อนศรี (2541 : บทคัดย่อ) ศึกษาความสัมพันธ์ของทักษะกระบวนการทางวิทยาศาสตร์กับความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ กลุ่มตัวอย่างเป็นนักเรียนมัธยมศึกษา ปีที่ 3 โรงเรียนสังกัดกรมสามัญศึกษาส่วนกลาง กลุ่มที่ 5 ปีการศึกษา 2540 จำนวน 552 คน โดยวิธีการสุ่มแบบสองขั้นตอน เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลประกอบด้วย แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ด้านการสังเกต การวัด การจำแนกประเภท การหาความสัมพันธ์ระหว่างสเปกกับสเปส และสเปกกับเวลา การคำนวณ การจัดกระทำและสื่อความหมายข้อมูล การลงความเห็นจากข้อมูล การพยากรณ์ การตั้งสมมุติฐาน การกำหนดนิยามเชิงปฏิบัติการ การกำหนดและควบคุมตัวแปร การทดลอง การตีความหมายข้อมูลและลงข้อสรุป และแบบทดสอบวัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ ผลการวิจัยพบว่า ทักษะกระบวนการทั้ง 13 ทักษะกับความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ มีความสัมพันธ์ทางบวกกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และทักษะกระบวนการทางวิทยาศาสตร์ด้านการจำแนกประเภท ส่งผลต่อความสามารถในการแก้ปัญหาทางวิทยาศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ส่วนทักษะกระบวนการทางวิทยาศาสตร์ด้านการสังเกต การจัดกระทำและสื่อความหมายข้อมูล และด้านการตีความหมายข้อมูลและลงข้อสรุป ส่งผลต่อความสามารถในการแก้ปัญหาทางวิทยาศาสตร์อย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนทักษะกระบวนการทางวิทยาศาสตร์ด้านอื่น ๆ ส่งผลต่อความสามารถในการแก้ปัญหาทางวิทยาศาสตร์อย่างไม่มีนัยสำคัญทางสถิติ

จิตฐิพร ศิริตานนท์ (2542 : บทคัดย่อ) ได้ศึกษา ผลการประเมินผลด้วยพหุโพลิโอที่มีต่อความรับผิดชอบ เจตคติต่อวิทยาศาสตร์ และทักษะกระบวนการทางวิทยาศาสตร์ กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 1 ปีการศึกษา 2540 โรงเรียนนวมินราชูทิศ โรงเรียนสตรีวิทยา 2 กรุงเทพมหานคร จำนวน 60 คน แบ่งเป็นกลุ่มทดลองและกลุ่มควบคุม พบว่า

1. ด้านเจตคติต่อวิทยาศาสตร์และความรับผิดชอบของกลุ่มทดลองและกลุ่มควบคุมแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ด้านทักษะกระบวนการทางวิทยาศาสตร์แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 โดยกลุ่มที่ประเมินผลด้วยพหุโพลิโอมีทักษะกระบวนการทางวิทยาศาสตร์สูงกว่ากลุ่มที่ประเมินผลแบบปกติ

5.3 งานวิจัยที่เกี่ยวกับการสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ในต่างประเทศ

งานวิจัยต่างประเทศที่ผู้วิจัยใช้ศึกษาเป็นแนวทางในการสร้างแบบทดสอบในครั้งนี้มีดังนี้

มอลิเตอร์ (Moliter. 1972 : 6241-A-6242-A) ได้พัฒนาแบบทดสอบประเมินผลทักษะการสรุปอ้างอิง (Inference) การทำนาย (Predictor) และการตรวจสอบ (Verification) โดยใช้ชื่อแบบทดสอบว่า Science Process Skills Test (SPST) แบบทดสอบฉบับนี้ใช้ทดสอบนักเรียนโดยการสัมภาษณ์เป็นรายบุคคล จากผลการวิจัยครั้งนี้ ทำให้ทราบแนวโน้มที่สามารถนำมาใช้ในการสร้างแบบทดสอบให้มีประสิทธิภาพได้ คือ ข้อสอบที่มีใจความเกี่ยวกับเรื่องที่นักเรียนคุ้นเคยมาก ๆ จะไม่สร้างแรงงู้อะไรเท่าที่ควร นักเรียนส่วนใหญ่มีประสบการณ์ในด้านการแสดงความคิดเห็นมากกว่าการตรวจสอบความคิดเห็น แบบทดสอบดังกล่าวสามารถใช้ประเมินผลนักเรียนได้เฉพาะส่วนที่เป็นการวัดทักษะการสรุปอ้างอิงและการตรวจสอบ ส่วนแบบทดสอบวัดทักษะการทำนายจำเป็นต้องได้รับการปรับปรุงแก้ไข เนื่องจากมีความเชื่อมั่นต่ำ

โรบินสัน (Robinson. 1973 : 1522-A) ได้พัฒนาแบบทดสอบวัดทักษะด้านการควบคุมตัวแปร (Controlling Variables) และการแปลความหมายจากข้อมูล (Interpreting Data) แบบทดสอบสองฉบับนี้ เป็นแบบปรนัยชนิดเลือกตอบหรือจากการได้ทำการวิเคราะห์เป็นรายข้อแล้ว ได้นำข้อสอบที่มีค่าอำนาจจำแนกตั้งแต่ .20 ขึ้นไป ค่าสัมประสิทธิ์สหสัมพันธ์กับคะแนนที่ได้จากการสอบเป็นรายบุคคล (Individual Test) ได้ค่าสหสัมพันธ์เป็น .66 แบบทดสอบอีกฉบับหนึ่ง

พีเรซ (Perez. 1978 : 3496-A) สร้างขึ้นสำหรับใช้ในประเทศฟิลิปปินส์ ลักษณะของแบบทดสอบประกอบด้วย ทักษะการสังเกต 11 ข้อ การเปรียบเทียบ 6 ข้อ การหาจำนวน 8 ข้อ การจำแนกประเภท 11 ข้อ การลงข้อสรุป 7 ข้อ การทำนาย 5 ข้อ และการทดลอง 11 ข้อ เมื่อคำนวณความเชื่อมั่นทั้งฉบับด้วยสูตร KR-20 ได้เท่ากับ .87 ค่าความยากง่ายเฉลี่ยเท่ากับ .39 และข้อสอบมีค่าอำนาจจำแนกที่มีประสิทธิภาพในการนำไปใช้ มีจำนวน 48 ข้อ เป็นแบบทดสอบที่พัฒนาขึ้นให้มีความสัมพันธ์กับเกณฑ์ภายนอกคือ The Individual Competency Measures อยู่ในระดับสูงและแบบทดสอบอีกฉบับหนึ่งมีความเชื่อมั่นสูง สามารถนำไปใช้กับเกณฑ์การท้าวิจัยเกี่ยวกับทักษะกระบวนการทางวิทยาศาสตร์

ดูเรน (Duren. 1978 : 24-25) กล่าวถึง Science Problem Solving Skill Test ของ Diet และ George ใช้สำหรับนักเรียนเกรด 1-3 ข้อสอบแต่ละข้อประกอบด้วย รูปภาพที่เหมาะสมกับสติปัญญาของเด็กแต่ละระดับ เกณฑ์ที่ใช้วัด คือ ความสามารถในการรวบรวมข้อมูลและความสามารถในการอธิบายข้อความที่เป็นเหตุเป็นผล

เบิร์น และคนอื่น ๆ (Burns and other. 1985 : 169) ได้พัฒนาแบบทดสอบชื่อว่า Development of a Integrated Process Skills Test : TIPS 2 โดยมีจุดมุ่งหมายเพื่อพัฒนาความเที่ยงตรง ความเชื่อมั่นของนักเรียนระดับชั้นมัธยมศึกษาตอนปลาย แบบทดสอบมีจำนวน 36 ข้อ หลังจากทดสอบกับกลุ่มตัวอย่างแล้วได้คะแนนเฉลี่ยของแบบทดสอบเท่ากับ 19.14 ค่าความเชื่อมั่นทั้งฉบับเท่ากับ .86 ความยากง่ายเท่ากับ .53 ค่าอำนาจจำแนกเฉลี่ยเท่ากับ .35 สัมประสิทธิ์สหสัมพันธ์ของแบบทดสอบแต่ละทักษะกับ TIPS 2 และ TIPS เป็น .86 และ .90 แบบทดสอบฉบับนี้ ทำให้มีข้อสอบร่วมในการวัดทักษะกระบวนการทางวิทยาศาสตร์เพิ่มขึ้น

เพาเวอร์ และโฟว์เลส (Powers and Fowles. 1998) ได้ทำการศึกษางานวิจัยเรื่อง ผลของการเปิดเผยหัวข้อของการสอบแบบความเรียงก่อนการสอบ เพื่อวัดผลของแบบวัดภาคปฏิบัติและความเที่ยงตรงของเครื่องมือ โดยการบอกหัวข้อของข้อสอบก่อนการทดสอบจริง ผู้ถูกวิจัยที่ตอบรับเข้าร่วมจำนวน 300 คน ทั้งหมดเป็นนักเรียนที่ใกล้จบการศึกษามาตอบข้อสอบ 2 ลักษณะ คือ 1 ตอบโดยไม่เคยรู้หัวข้อข้อสอบมาก่อน และอีก 1 ตอบโดยรู้หัวข้อของข้อสอบก่อนสอบ 2 สัปดาห์ รวมแล้วแต่ละคนจะต้องข้อสอบ 4 ข้อ การวิจัยนี้ผลการวิเคราะห์ไม่ปรากฏความแตกต่างอย่างมีนัยสำคัญทางสถิติ ระหว่างผลการเปิดเผยหัวข้อก่อนปฏิบัติกับการปฏิบัติโดยมีทักษะการเขียน ข้อจำกัดของการศึกษาถูกอภิปรายในภาพของกลุ่มทฤษฎีสรุปอ้างอิง ที่มีสถานการณ์เสี่ยงสูงและขอบเขตที่จะให้ค่าที่ได้เป็นบวกและลบได้

การ์เนอร์ และ เอนเกลฮาร์ด (Garner and Engelhard. 1999) ได้ศึกษางานวิจัยเรื่อง ความแตกต่างของเพศในการสอบปฏิบัติโดยใช้แบบทดสอบปรนัย และแบบทดสอบอัตนัย วิชาคณิตศาสตร์ งานวิจัยนี้มุ่งศึกษาผลของความแตกต่างของเพศในการสอบปฏิบัติที่ใช้แบบทดสอบปรนัยและอัตนัยวิชาคณิตศาสตร์ระดับมัธยมศึกษาตอนปลาย โดยสร้างขึ้นตามหลักสูตรที่กำหนดขึ้น จุดมุ่งหมายที่สองเพื่อแสดงบทพิสูจน์วิธีการแบบฟาเซทแบบราซท์โมเดล (1980) สำหรับเปรียบเทียบการปฏิบัติที่มีรูปแบบแตกต่างกันและค้นหารูปแบบที่มีฟังก์ชันต่างกัน (DIF) กลุ่มตัวอย่างที่ใช้เป็นนักเรียนระดับ 11 จำนวน 3,952 คน ที่เข้าเรียนในปี 1994 และผ่านการทดสอบด้วยแบบทดสอบ Georgia High School Graduation Test ในการวิเคราะห์ผลจากทั้งสองกลุ่มได้ค่าเฉลี่ยและค่า DIF ของเพศหญิง แสดงความแตกต่างอย่างมีนัยสำคัญ และมีความคงที่ภายในสูงกว่านักเรียนชาย เมื่อใช้แบบทดสอบแบบปรนัยวิชาพีชคณิตและการวัด , ตัวเลขและการคำนวณ , การวิเคราะห์ข้อมูล และเหตุผลเฉพาะ คะแนนเฉลี่ยของเพศชายให้ค่าแตกต่างอย่างมีนัยสำคัญสูงกว่าเพศหญิง 2 ใน 8 ข้อของข้อสอบแบบอัตนัย อย่างไรก็ตาม เมื่อเพศชายและเพศหญิงก็ให้ค่าสอดคล้องกัน ส่วนด้านความสามารถ ให้ค่าแตกต่างกันอย่างมีนัยสำคัญเมื่อปฏิบัติโดยใช้แบบทดสอบอัตนัย สรุปได้ว่า ความแตกต่างของเพศในการทดสอบคณิตศาสตร์มีคุณภาพดีหรือไม่ขึ้นอยู่กับเนื้อหาและรูปแบบของข้อสอบ

5.4 งานวิจัยที่เกี่ยวกับการสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ในประเทศ

เสริม ทัดศรี (2522 : 23-43) สร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์เพื่อต้องการวิเคราะห์องค์ประกอบความสามารถทางการเรียนของนักเรียนชั้นมัธยมศึกษาปีที่ 1 จังหวัดสงขลา ลักษณะข้อสอบเป็นแบบปรนัยชนิดเลือกตอบ 5 ตัวเลือก จำนวน 96 ข้อ วัดความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์ 12 ด้าน คือ การสังเกต การจำแนก การวัด การอธิบาย การตั้งปัญหา การพยากรณ์ การตั้งสมมุติฐาน การทดลอง การตีความข้อมูล การสรุป การนำไปใช้ และการฝึกทักษะ คำนวณค่าความเชื่อมั่นด้วยสูตร KR-20 ได้ค่าความเชื่อมั่นแต่ละทักษะดังนี้ .3986 , .3254 , .4369 , .3204 , .4237 , .3238 , .3277 , .3763 , .3743 , .3552 , .4558 , .3164 ตามลำดับ และค่าความเชื่อมั่นทั้งฉบับ .8753 สรุปว่า ทักษะเหล่านี้จะมี 5 องค์ประกอบ องค์ประกอบที่ 1 ประกอบด้วย ความสามารถในการสังเกต องค์ประกอบที่ 2 ประกอบด้วย ความสามารถในการตีความและการสรุป องค์ประกอบที่ 3 ประกอบด้วย ความสามารถในการจำแนกและการตั้งปัญหา องค์ประกอบที่ 4 ประกอบด้วย ความสามารถในการพยากรณ์ และการตั้งสมมุติฐานและการนำไปใช้ องค์ประกอบที่ 5 ประกอบด้วย ความสามารถในการวัด การอธิบาย การทดลอง และการฝึกทักษะ

อนันต์ จันทกรวี (2525 : 48) สร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นผสม (Integrated Skill) เพื่อใช้ศึกษาเรื่อง ผลของการใช้คำถามของครูที่มีต่อทักษะกระบวนการทางวิทยาศาสตร์ ผลสัมฤทธิ์และทัศนคติของนักเรียนชั้นมัธยมศึกษาปีที่ 2 (ม.ศ. 2 เดิม และ ม. 2 ใหม่) ลักษณะข้อสอบเป็นแบบปรนัย ชนิดเลือกตอบ 5 ตัวเลือก ประกอบด้วย ทักษะการจัดกระทำข้อมูล การแปลความหมายข้อมูล และการสรุป การสร้างสมมุติฐาน การออกแบบการทดลองและดำเนินการทดลอง จำนวน 28 ข้อ คำนวณค่าความเชื่อมั่นด้วยสูตร KR-21 ได้ความเชื่อมั่นเท่ากับ .3760 , .2941 , .4949 ตามลำดับ และค่าความเชื่อมั่นทั้งฉบับ .6695 เมื่อคำนวณด้วยวิธี Split-half จะมีค่าความเชื่อมั่นแต่ละทักษะเป็น .3966 , .4674 , .4109 , .5453 ตามลำดับ และค่าความเชื่อมั่นทั้งฉบับ .7536

ก่อศักดิ์ ศรีน้อย (2527 : 51) สร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์เพื่อทำการวิจัยเรื่อง การใช้คำถามเน้นทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานและขั้นบูรณาการในการสอนแบบสืบเสาะหาความรู้ที่มีผลต่อความคิดสร้างสรรค์ทางวิทยาศาสตร์และผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ลักษณะข้อสอบเป็นแบบตอบสั้น โดยยึดตามแผนการสอนเรื่อง การขนส่ง และการสื่อสาร แบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ เป็นแบบทดสอบชนิดเลือกตอบ 5 ตัวเลือก จำนวน 50 ข้อ คำนวณค่าความเชื่อมั่นด้วยสูตร KR-20 ได้ค่าความเชื่อมั่นเท่ากับ .8211

รวีวรรณ อังคนุรักษ์พันธุ์ (2531 : บทคัดย่อ) ได้สร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 โรงเรียนสาธิต สังกัดทบวงมหาวิทยาลัยของรัฐ จำนวน 1,095 คน ประกอบด้วยข้อคำถามที่วัดองค์ประกอบที่สำคัญของทักษะกระบวนการทางวิทยาศาสตร์ 13 ทักษะ คือ การสังเกต การวัด การจำแนกประเภท การหาความสัมพันธ์ระหว่างสเปกกับสเปกและสเปกกับเวลา การคำนวณ การจัดกระทำและสื่อความหมายข้อมูล การลงความคิดเห็นจากข้อมูล การพยากรณ์ การทดลอง การตั้งสมมุติฐาน การกำหนดนิยามเชิงปฏิบัติการ การกำหนดและการควบคุมตัวแปร การตีความหมาย ข้อมูลและลงข้อสรุป ผลการศึกษาพบว่า ค่าความยากและค่าอำนาจจำแนกของแบบทดสอบมีค่าอยู่ในช่วง .21 ถึง .78 และ .20 ถึง .85 ตามลำดับ คำนวณค่าความเชื่อมั่นด้วยสูตร KR-20 ได้ค่าความเชื่อมั่นของแบบทดสอบแต่ละทักษะมีค่าอยู่ในช่วง .2290 ถึง .5769 และรวมทั้งฉบับมีค่าเท่ากับ .8383 มีนัยสำคัญทางสถิติที่ระดับ .01 ค่าความเที่ยงตรงเชิงโครงสร้างมีค่าน้ำหนักองค์ประกอบอยู่ในช่วง .1972 ถึง .6202 มีนัยสำคัญทางสถิติที่ระดับ .01 และเกณฑ์ปกติของแบบทดสอบอยู่ในช่วง T15 ถึง T87

วัฒนชัย จันทรวินกุล (2538 : บทคัดย่อ) ได้สร้างแบบทดสอบภาคปฏิบัติวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐาน กลุ่มตัวอย่างเป็นนักเรียนชั้นประถมศึกษาปีที่ 5 จังหวัดปทุมธานี ปีการศึกษา 2537 จำนวน 154 คน ผลการศึกษาพบว่า ค่าความยากและอำนาจจำแนกของแบบทดสอบอยู่ในช่วง .21 ถึง .70 และ .23 ถึง .67 ตามลำดับ ความเชื่อมั่นของแบบทดสอบคำนวณโดยใช้สูตรสัมประสิทธิ์แอลฟา ในแต่ละทักษะอยู่ในช่วง .436 ถึง .570 และรวมทั้งฉบับมีค่าเท่ากับ .750 ค่าความเที่ยงตรงเชิงโครงสร้างของแบบทดสอบคำนวณโดยวิธี Multitrait Multimethod มีความเที่ยงตรงเชิงเหมือนอยู่ในช่วง .559 ถึง .675 ความเที่ยงตรงเชิงจำแนกอยู่ในช่วง -.241 ถึง .542 สัมประสิทธิ์สหสัมพันธ์ภายในแต่ละทักษะและทั้งฉบับคำนวณโดยใช้สูตรของเพียร์สัน อยู่ในช่วง .268 ถึง .807 และ .577 ถึง .891 ตามลำดับ เป็นความเที่ยงตรงเชิงโครงสร้างอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และมีเกณฑ์ปกติของแบบทดสอบทั้งฉบับและแต่ละทักษะ

5.5 งานวิจัยที่เกี่ยวข้องกับความเชื่อมั่นของวิธีการประเมินและผู้ประเมินในต่างประเทศ

สตีเฟ่น (Steven, 1991) ได้ศึกษาเกี่ยวกับความเชื่อมั่นของผู้ตรวจ เรื่อง ผู้ตัดสินตัวอย่างงานเขียนของวิทยาลัยโกลเดนเวสต์ : การประเมินความเชื่อมั่นของผู้ตรวจ จุดประสงค์ต้องการศึกษาตัวอย่างงานเขียนของนักเรียน โดยให้คะแนนแบบประเมินภาพรวม (Holistic) โดยใช้ผู้ตรวจเป็นรายคู่ ด้วยเกณฑ์การวัดแบบ 6 ระดับ (Six-point Scale) ผู้ตรวจต้องทำเครื่องหมาย + หรือ - วิธีนี้เป็นการเปลี่ยนจากการให้คะแนนแบบจำนวนเต็มมาเป็นแบบทศนิยม ในปี 1991 ตัวอย่างงานเขียนจำนวน 499 เรื่อง ที่นำมาจากส่วนหนึ่งของการทดสอบเพื่อจัดอันดับสำหรับนักเรียนในโครงการ SOAR ของ Coast Community College District (เขตวิทยาลัยชุมชนแถบชายฝั่ง) ที่วิเคราะห์การให้คะแนนที่น่าเชื่อถือของ ผู้ตรวจ งานวิจัยนี้ใช้ 2 วิธี เพื่อประเมินความคงที่ในการตัดสินตัวอย่างงานเขียน วิธีแรกที่ใช้เป็นการคำนวณความแตกต่างระหว่างค่าบ่งบอกระดับตัวอย่างงานเขียน 2 ค่า มีค่าเท่ากับ 34.4 % ของคู่ที่ตัดสินทั้งหมด ค่าที่บ่งบอกระดับจากผู้ตัดสินแต่ละคนจะเหมือนกัน แต่ละค่าจะแตกต่างกันไป 1 ใน 3 ระดับ มี 30.9 % ของตัวอย่าง และ 2 ใน 3 ระดับ มี 15.4 % ของตัวอย่าง และ 1 ระดับ มี 13.4 % โดยรวมแล้วค่าบ่งบอกระดับจะแตกต่างออกไป 1 ระดับ หรือน้อยกว่า 94.1 % ของตัวอย่าง ซึ่งเกินมาตรฐานต่ำสุดที่ 90 % ที่แนะนำโดย California Community Colleges (CCC) วิธีที่สองที่ใช้ก็คือ สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน แสดงความสัมพันธ์ระหว่างการตรวจเป็นคู่ สำหรับตัวอย่างข้อเขียน ค่าความสัมพันธ์มีค่าเป็นบวก ($r = .76$) ซึ่งเกินมาตรฐานต่ำสุดที่ .75 ตามคำแนะนำของ CCC และข้อมูลที่ได้อบรมรับสมมุติฐาน ที่ว่าการตรวจจากผู้ตรวจแต่ละคนได้ผลสอดคล้องกัน

เทสส์, แมร์ และคณะ (Tasse, Mare J. and others. 1994) ศึกษาความสอดคล้องของผู้ตรวจ เรื่อง แบบทดสอบซ้ำ และความเชื่อมั่นระหว่างผู้ตรวจแบบมาตรวัดพฤติกรรมการปรับตัวของคิวเบก (Quebec) ซึ่งได้รับการนำไปใช้อย่างแพร่หลายในเมืองคิวเบก (ประเทศแคนาดา) เพื่อใช้ประเมินพฤติกรรมของผู้ที่มีระดับสติปัญญาต่ำในสภาพแวดล้อมเกี่ยวกับการศึกษา การทำงาน การอยู่อาศัย หรือ สภาพแวดล้อมในโรงพยาบาล การศึกษานี้ประเมินความเชื่อมั่นของผู้ตรวจและความเชื่อมั่นโดยวิธีสอบซ้ำของ QABS และเพื่อวัดความเชื่อมั่นของการสอบซ้ำ QABS จึงได้รับการประเมินโดยผู้ตรวจจำนวน 27 คน กับกลุ่มตัวอย่างจำนวน 45 คน (อายุระหว่าง 6 ถึง 47 ปี) จากบ้านเรือน สถานที่พักอาศัย โรงเรียน และโรงพยาบาลต่าง ๆ และเพื่อให้ได้ค่าความเชื่อมั่นของผู้ตรวจ ได้ใช้กลุ่มตัวอย่างอีกกลุ่มจำนวน 44 คน (อายุระหว่าง 3 ถึง 54 ปี) แต่ละคนจะได้รับการประเมินจากผู้ตรวจ 2 คน ค่าสัมประสิทธิ์ความเชื่อมั่นระหว่างการตรวจโดยใช้วิธีประเมินภาพรวม (Global Score) กับระดับพฤติกรรมการปรับตัว มีค่าเท่ากับ .92 คำนวณโดยวิธีสอบซ้ำ และ ระหว่างผลรวมของคะแนนดิบ มีค่าเท่ากับ .99 ค่าสัมประสิทธิ์ความสอดคล้องของผู้ตรวจที่ใช้การตรวจแบบประเมินภาพรวม (Global Score) กับระดับพฤติกรรมการปรับตัว มีค่าเท่ากับ .83 และ ระหว่างผลรวมของคะแนนดิบ เท่ากับ .99 พบว่า QABS เป็นมาตรวัดพฤติกรรมการปรับตัวแบบอิงเกณฑ์ที่มีความเชื่อถือได้สูง มีเกณฑ์ปกติ และแบบเกณฑ์การให้คะแนนที่ชัดเจนง่ายต่อการใช้งาน

ไครท และ สตีลเชอร์ (Klein and Stecher. 1998) ศึกษาเกี่ยวกับวิธีการตรวจให้คะแนน เรื่อง การตรวจแบบวิเคราะห์ย่อยกับแบบประเมินรวมสำหรับการทดสอบภาคปฏิบัติวิชาวิทยาศาสตร์ งานวิจัยนี้ต้องการศึกษา 2 ส่วน คือ หาความสอดคล้องภายในของผู้ตรวจ (Interreader) , ความเชื่อมั่นของการตรวจให้คะแนน (Score reliability) และความต้องการเวลาในการตรวจของผู้ตรวจจำนวน 3 คน สำหรับการ

สอบภาคปฏิบัติวิชาวิทยาศาสตร์ การศึกษาแรกเกี่ยวกับการให้คะแนนคำตอบที่ได้จากนักเรียนเกรด 5 , 8 และ 10 ใน 3 จุดมุ่งหมาย ("หลักสูตรมาตรฐาน") โดยวิธีการปฏิบัติ การศึกษาต่อมาคือต้องการคำนวณคะแนนทั้ง 3 ระดับที่ได้มา การศึกษาทั้ง 2 ข้อแรกใช้การตรวจแบบวิเคราะห์ย่อยและการประเมินรวม เพื่อให้ระดับคำตอบที่ได้ แตกต่างกันในด้านคุณลักษณะย่อยของรูปรีคเท่านั้น การตรวจแบบวิเคราะห์ย่อยจะใช้เวลามากแต่ก็ยังให้ค่าความสอดคล้องภายในของผู้ตรวจสูง แต่กระนั้นเมื่อทำการเฉลี่ยจากทุก ๆ คำถามของเครื่องมือแล้ว คะแนนของนักเรียนที่ตรวจแบบประเมินรวมให้ค่าความเชื่อมั่นเหมือนกับคะแนนที่ได้จากการตรวจแบบวิเคราะห์ย่อย แสดงว่าทั้งสองวิธีมีความสัมพันธ์กันสูงเมื่อลดความสอดคล้องภายในระหว่างผู้ตรวจ โดยใช้ผู้ตรวจ 2 คน ต่อ 1 คำถาม จะไม่ปรากฏผลต่อค่าที่คำนวณได้

ปอมพลุน และ ซันบาย (Pomplun and Sundbye. 1999) ได้ศึกษาความแตกต่างของเพศ ของแบบทดสอบแบบขยายความ วิชาการอ่าน (Constructed Response Reading Items) การศึกษาครั้งนี้ต้องการค้นหาความแตกต่างของเพศในการตอบแบบทดสอบการอ่านจากโครงการประเมินของรัฐบาล แยกออกเป็นองค์ประกอบที่สัมพันธ์เชิงโครงสร้างและไม่สัมพันธ์กัน จัดให้เหมือนกันโดยผู้เชี่ยวชาญทางเนื้อหาวิชา ที่เกี่ยวกับชีววรรณคดี และวิชานำร่อง (A pilot study) ใช้ผู้ตรวจ 4 คน ที่ได้รับการฝึกอบรมสำหรับการให้คะแนน ใช้กระดาษคำตอบจำนวน 500 แผ่น แบ่งออกเป็น 2 ระดับ และแบ่งเป็นระหว่างเพศประกอบด้วย ส่วนคำตอบ (เฉลย) , ส่วนคำตอบที่ไม่สัมพันธ์กัน , คำตอบผิด , จำนวนคำที่เขียน และวิธีการวัดที่สับสน แบบฟอร์มการให้คะแนนรูปรีคของรัฐบาล ความสัมพันธ์ระหว่างสิ่งต่าง ๆ ที่ต้องการศึกษาเป็นการตรวจแบบประเมินรวม (Holistic Score) , การให้คะแนนย่อย , ความแตกต่างของเพศที่ต้องการค้นหาความสัมพันธ์ ความคิดและการวิเคราะห์แบบถดถอย ผลที่ได้คือ จำนวนคำและจำนวนคำตอบที่ไม่สัมพันธ์กัน กับความแตกต่างระหว่างเพศ เมื่อใช้วิธีการตรวจแบบประเมินรวม แสดงค่าทางสถิติแตกต่างกันอย่างมีนัยสำคัญ และค่าทางสถิติแตกต่างกันอย่างมีนัยสำคัญ เมื่อศึกษาตัวแปรควบคุม ซึ่งงานนี้ต้องการสืบค้นสิ่งที่ส่งผลกระทบต่อความแตกต่างของผู้ตรวจที่มีต่อนักเรียนชายและนักเรียนหญิง

5.6 งานวิจัยที่เกี่ยวข้องกับความเชื่อมั่นของวิธีการประเมินและผู้ประเมินในประเทศ

ดำเกิง วงศ์สว่าง (2528) ได้ศึกษาผลของการทดสอบย่อยโดยใช้แบบทดสอบอัตนัย เดิมคำ และเลือกตอบ ที่มีต่อความสามารถในการเขียนภาษาไทย ของนักเรียนชั้นประถมศึกษาปีที่ 3 โรงเรียนสามเสนนอก จำนวน 144 คน ผลการวิจัยพบว่า ในส่วนการตรวจให้คะแนนแบบทดสอบวัดความสามารถในการเขียนภาษาไทย 1 ฉบับ จำนวน 3 ข้อ ของผู้ตรวจ 2 คน โดยกำหนดเกณฑ์การให้คะแนนยึดตามแนวของประสิทธิ์ กาพย์กลอน นั้น ข้อสอบข้อ 1 ข้อ 2 และข้อ 3 มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.8682 0.8849 และ 0.9454 ตามลำดับ และเมื่อทดสอบนัยสำคัญ พบว่า มีนัยสำคัญทางสถิติที่ระดับ .01 ซึ่งแสดงว่าการตรวจให้คะแนนของผู้ตรวจ 2 คน มีความเชื่อถือได้

จิราพร อัครสมพงษ์ (2536) ได้ศึกษาเปรียบเทียบความสอดคล้องของคะแนนแบบสอบความเรียงจากผู้ตรวจให้คะแนนและวิธีการตรวจต่างกัน ที่มีวัตถุประสงค์เพื่อ ศึกษาเปรียบเทียบความสอดคล้องในการให้คะแนนของผู้ตรวจให้คะแนน 2 คน ที่ตรวจแบบสอบความเรียงโดยใช้วิธีเดียวกัน และความสอดคล้องในการให้คะแนนระหว่างวิธีการตรวจที่แตกต่าง พบว่าความสัมพันธ์ในการให้คะแนนระหว่างผู้ตรวจให้คะแนนด้วยวิธีเดียวกัน ผู้ตรวจให้คะแนนที่ตรวจด้วยวิธีเดียวกันมีความสัมพันธ์กันสอดคล้องกัน ในทิศทางเดียวกัน

และมีค่าสัมประสิทธิ์สหสัมพันธ์ของการตรวจแบบวิธีวิเคราะห์ให้เสร็จเป็นรายข้อกับการตรวจแบบวิธีประเมินรวมให้เสร็จเป็นรายข้อมีค่าสูงสุด คือ .8116

สมพงษ์ อัครสุภฤกษ์ (2542 : บทคัดย่อ) ศึกษาเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่มีการตรวจและความยาวของแบบทดสอบต่างกัน ด้วยแบบทดสอบเลือกตอบถูกผิด วิชาสถิติธุรกิจ 1 จำนวน 2 ฉบับ และนำมาตรวจด้วยให้คะแนนด้วยวิธีต่างกัน กลุ่มตัวอย่างเป็นนักศึกษาประกาศนียบัตรวิชาชีพชั้นสูง สถาบันเทคโนโลยีราชมงคล ส่วนกลาง จำนวน 728 คน พบว่า

1. ค่าความเชื่อมั่นของแบบทดสอบเลือกตอบถูกผิดที่มีจำนวนข้อต่างกันและตรวจให้คะแนนด้วยวิธีเดียวกันคือสูตรสัมประสิทธิ์แอลฟาและสูตรของราซ จะมีความแตกต่างอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

2. ค่าความเชื่อมั่นของแบบทดสอบเลือกตอบถูกผิดที่มีการตรวจให้คะแนนต่างกัน และจำนวนข้อเท่ากันด้วยสูตรสัมประสิทธิ์ของราซ จะมีความแตกต่างอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01

พรรณี เจียมสุขบุตร (2543) ได้ศึกษาเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจและวิธีการตรวจต่างกัน การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจ 2 คน และ 3 คน โดยใช้วิธีการตรวจแบบวิธีรวม และวิธีการวิเคราะห์ โดยใช้กับกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2542 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดกรุงเทพฯ กลุ่ม 7 จำนวน 70 คน ซึ่งเลือกมาโดยการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการศึกษาเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 เรื่อง สมการและอสมการ ผลการศึกษาได้ผลดังนี้

1. ค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ โดยการตรวจให้คะแนนด้วยวิธีการวิเคราะห์ที่มีจำนวนผู้ตรวจ 2 คน และ 3 คน มีค่าเท่ากับ 0.92 และ 0.96 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนการตรวจให้คะแนนด้วยวิธีรวม ที่มีจำนวนผู้ตรวจ 2 และ 3 คน มีค่าเท่ากับ 0.97 และ 0.98 ตามลำดับ ซึ่งแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

2. ค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ที่ใช้วิธีการตรวจให้คะแนนแบบวิธีรวม และวิธีวิเคราะห์ของผู้ตรวจให้คะแนน 2 คน มีค่าเท่ากับ 0.97 และ 0.92 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนแบบทดสอบที่ใช้วิธีการตรวจให้คะแนนแบบวิธีรวม และวิธีการวิเคราะห์ของผู้ตรวจให้คะแนน 3 คน มีค่าเท่ากับ 0.98 และ 0.96 ตามลำดับแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

5.7 งานวิจัยที่เกี่ยวข้องกับความเชื่อมั่นของการให้คะแนนตามทฤษฎีการสรุปอ้างอิงในต่างประเทศ

สมิท (Smith. 1978 : 319-346) ได้ศึกษาผลของความคลาดเคลื่อนในการสุ่มกลุ่มตัวอย่าง เพื่อนำมาใช้ในการศึกษาตามโครงสร้างทฤษฎีการสรุปอ้างอิง ในลักษณะหลายองค์ประกอบ (Multifacet Study) ที่ต้องการใช้กลุ่มตัวอย่างจำนวนน้อย ๆ และเสนอแนะวิธีการลดความคลาดเคลื่อนเชิงสุ่มไว้ 3 วิธี ดังนี้

1. สุ่มระดับขององค์ประกอบแต่ละตัว ที่ต้องการศึกษาในแบบจำลองการวิเคราะห์ความแปรปรวนให้ได้จำนวนมากที่สุดเท่าที่จะทำได้

2. สุ่มระดับของผลกระทบแต่ละตัวให้เป็นสัดส่วนกับประชากรที่ศึกษา

3. เปลี่ยนรูปแบบของความสัมพันธ์ระหว่างองค์ประกอบ เพื่อลดจำนวนของแหล่งความแปรปรวน

กรณีนี้ศึกษาความแปรปรวนของประชากร และจำนวนของกลุ่มตัวอย่างที่มีจำนวนมากภายใต้ข้อจำกัด ในกรณีนี้ใช้วิธีลดความคลาดเคลื่อนในข้อ 2 และ 3 ก็เพียงพอ แต่ถ้านการสุ่มระดับขององค์ประกอบมีจำนวนน้อย ไม่ว่าจะใช้วิธีใดก็ตาม ก็ไม่สามารถลดความคลาดเคลื่อนลงได้

มาครีเอตี้ (Macready. 1983 : 149-157) ศึกษาการใช้ทฤษฎีการสุ่มอย่างอิงในการวิเคราะห์การสอบ เพื่อประเมินความสัมพันธ์ระหว่างข้อสอบ วัดคุณสมบัติของการศึกษาเพื่อประเมินการจัดกลุ่มข้อสอบอิงโดเมน (Domain) ที่สร้างขึ้น วิธีการที่ใช้เป็นการสำรวจกลุ่มของข้อสอบ ชุดที่มีความเป็นเอกพันธ์ซึ่งมีระดับความยากเท่ากัน องค์ประกอบที่ศึกษาประกอบด้วย โดเมนการคูณจำนวนเต็ม (d) ห้องเรียน (c) จำนวนหลักของตัวคูณ (n) เครื่องมือที่ใช้ในการศึกษา เป็นข้อสอบอิงโดเมนวิชาคณิตศาสตร์เนื้อหา เกี่ยวกับการคูณ โดยได้จัดกลุ่มข้อสอบเป็น 20 โดเมน แต่ในการวิจัยนี้ใช้ข้อสอบเพียง 6 โดเมน แต่ละโดเมนเลือกข้อสอบมาใช้ 10 ข้อ กลุ่มตัวอย่างเป็นนักเรียนเกรด 5 จำนวน 260 คน ผลการศึกษาพบว่า นักเรียนบางคนในห้องเรียนทำข้อสอบบางข้อในโดเมนข้อสอบทั้งหมด ($S : C \times d$) เป็นแหล่งความแปรปรวนที่มีค่ามากที่สุด ซึ่งสามารถอธิบายแหล่งความแปรปรวนทั้งหมดร้อยละ 47 จากผลที่ได้แสดงว่า โดเมนใดโดเมนหนึ่ง หรือหลาย ๆ โดเมนบรรจุข้อสอบที่ไม่เป็นเอกพันธ์ เมื่อพิจารณาผลการสอบของนักเรียน พบว่า มีการเบี่ยงเบนออกจากคุณสมบัติด้านความเป็นเอกพันธ์ของข้อสอบ แหล่งความแปรปรวนที่เหลืออีก 4 แหล่ง อธิบายความแปรปรวนทั้งหมดได้เพียงร้อยละ 1 หรือมากกว่า คือ $c : s : c : d$ และ $s : c \times d$ ซึ่งนำมารวมกันแล้ว อธิบายความแปรปรวนทั้งหมดได้ร้อยละ 51 เฉพาะแหล่งความแปรปรวนสุดท้าย หมายถึง ความยากของข้อสอบในแต่ละโดเมน สำหรับนักเรียนแต่ละคนมีค่าไม่เท่ากัน สำหรับแหล่งความแปรปรวนอื่น ๆ ที่เหลือมีค่าน้อยมาก โดยเฉพาะจำนวนหลักของตัวคูณ สรุปว่าไม่ควรใช้จำนวนหลักของตัวคูณ 4 ตำแหน่ง ใช้เพียง 3 ตำแหน่งก็เพียงพอ เพราะให้ค่าความยากของ ข้อสอบไม่แตกต่างกัน จากการประมาณค่าสัมประสิทธิ์การสุ่มอย่างอิงของข้อสอบแต่ละโดเมน โดยใช้รูปแบบการวิเคราะห์ 2 รูปแบบ คือ $s \times i$ และ $s \times r$ เมื่อ หมายถึงการสอบซ้ำ ค่าสัมประสิทธิ์การสุ่มอย่างอิงของข้อสอบหนึ่งข้อมีค่าอยู่ระหว่าง .3380 ถึง .6060

อิบ ราฮิม (Ibrahim. 1984) ได้ประยุกต์ทฤษฎีการสุ่มอย่างอิง เพื่อประมาณค่าความแปรปรวนที่มีต่อการประเมินวัดคุณสมบัติทางการศึกษา โดยสุ่มตัวอย่างครู 80 คน และนักศึกษา 80 คน ในประเทศซูดาน ประเมินวัดคุณสมบัติทางการศึกษา 2 ชนิด คือ วัดคุณสมบัติที่สำคัญจริง ๆ และวัดคุณสมบัติที่คาดหวัง ตัวประกอบในการศึกษามี ผู้ประเมิน กลุ่มผู้ประเมิน จำนวนครั้งของการประเมิน ถิ่นที่อยู่ของผู้ประเมิน ชนิดของวัดคุณสมบัติ สถานที่ทำงานของผู้ประเมิน และเพศของผู้ประเมิน พบว่า ตัวประกอบที่มีผลต่อการประเมินมากที่สุด คือ ผู้ประเมินและกลุ่มผู้ประเมิน ส่วนสถานที่ทำงานของผู้ประเมินมีผลเล็กน้อยเท่านั้น ส่วนตัวประกอบที่เหลือไม่มีผลต่อการประเมิน

โอ ไบรอน และ โจนส์ (O'Brien, R.M. and Jones, B. 1986 : 21-27) ศึกษาการใช้ทฤษฎีการสุ่มอย่างอิงเพื่อประมาณค่าความเชื่อมั่นของตัวแปรระดับโรงเรียน 16 ตัว เป็นค่าเฉลี่ยหรือร้อยละของกลุ่มตัวอย่างนักเรียน ตัวแปรที่ศึกษาแบ่งเป็น 5 กลุ่มคือ ค่าเฉลี่ยของสถานะทางสังคมและเศรษฐกิจของครอบครัว ค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียน ร้อยละของนักเรียนที่อาศัยอยู่กับบิดามารดา ร้อยละของคุณภาพของการบริการห้องสมุด ระบบการเรียนการสอนในโรงเรียน และร้อยละของนักเรียนที่ปฏิบัติตามกฎของโรงเรียน กลุ่มตัวอย่างเป็นนักเรียนเกรด 2 จำนวน 1122 โรงเรียนละ 36 คน ดำเนินการวิจัยโดยให้กลุ่มตัวอย่างตอบคำถามในแบบสอบถามของตัวแปรแต่ละตัว พบว่า ความเชื่อมั่นของแบบสอบถามเพิ่มขึ้นตามจำนวนผู้ตอบที่เพิ่มขึ้นจากแต่ละโรงเรียน จำนวนข้อคำถาม และจำนวนโรงเรียน ผลการวิจัยนี้แสดงให้เห็น

เห็นว่า จำนวนผู้ประเมิน จำนวนข้อคำถามของแบบสอบถามและจำนวนสถานที่ของผู้ประเมินมีผลต่อค่าความเชื่อมั่นของการประเมิน

คุนนอน (Kunnon. 1992 : 30 - 47) ได้ทำการวิจัยเรื่องการหาคุณภาพของแบบทดสอบอิงเกณฑ์ โดยใช้ทฤษฎีการสรุปอ้างอิง (G – Theory) การวิเคราะห์องค์ประกอบและการวิเคราะห์กลุ่ม (Factor and Cluster Analysis) การศึกษาวิจัยนี้ใช้วิธีการวิเคราะห์ 3 ชนิด คือ ทฤษฎีการสรุปอ้างอิง เพื่อตรวจสอบความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ ส่วนวิธีการวิเคราะห์องค์ประกอบและการวิเคราะห์กลุ่ม เพื่อตรวจสอบความเที่ยงตรงของแบบทดสอบอิงเกณฑ์ แบบทดสอบอิงเกณฑ์ที่ใช้ศึกษาคือ ESLPE (NESL – PE) ของมหาวิทยาลัยแคลิฟอร์เนีย โดยแบ่งข้อสอบเป็น 3 ฉบับคือ การฟัง การอ่านและศัพท์ ไวยากรณ์ ฉบับละ 30 ข้อเท่า ๆ กันทุกฉบับเป็นชนิดเลือกตอบ และแบ่งกลุ่มตัวอย่างออกเป็น 4 กลุ่มตามความสามารถในการสอบ การคำนวณหาความเชื่อมั่นโดยทฤษฎีสรุปอ้างอิง ใช้โปรแกรมสำเร็จรูป GENOVA ซึ่งแบ่งออกเป็น $p \times (i : t)$ คือ ข้อสอบที่แบ่งอยู่ในแบบทดสอบซึ่งเป็นโมเดลผสม (Mix Model) และการคำนวณค่าความเที่ยงตรง โดยใช้การวิเคราะห์องค์ประกอบและการวิเคราะห์กลุ่มใช้โปรแกรมสำเร็จรูป SPSS – X โดยแบ่งข้อสอบแต่ละฉบับออกเป็นอีก 2 ฉบับย่อย พบว่า ความเชื่อมั่นของแบบทดสอบแตกต่างกัน ทุก ๆ กลุ่มและค่าสูงในกลุ่มที่มีความสามารถต่ำ ในการหาความเชื่อมั่นโดยใช้คะแนนมวลความรู้ค่าความเชื่อมั่นของคะแนนรวมทั้งสี่กลุ่มจะสูงที่สุด ส่วนความเที่ยงตรงของแบบทดสอบโดยวิธีวิเคราะห์องค์ประกอบ พบว่า การแบ่งข้อสอบออกเป็น 2 ส่วนย่อย ๆ ในแต่ละองค์ประกอบ จะเป็นวิธีที่ดีที่สุดสำหรับกลุ่มต่าง ๆ และวัดทิศทางความสามารถ ในกลุ่มที่มีความสามารถต่ำ ได้ดีกว่าวัดในระดับที่สูงกว่า และการวิเคราะห์กลุ่ม (Cluster Analysis) พบว่านักเรียนจะได้รับการตัดสินแตกต่างไปจากเดิม ถ้าใช้คะแนนในส่วนย่อยเป็นตัวตัดสิน

5.8 งานวิจัยที่เกี่ยวข้องกับความเชื่อมั่นของการให้คะแนนตามทฤษฎีการสรุปอ้างอิงในประเทศ

งานวิจัยภายในประเทศเกี่ยวกับ ความเชื่อมั่นของผู้ตรวจให้คะแนน นั้นยังพบน้อยมาก ที่ผู้วิจัยได้รวบรวมได้ มีดังนี้

แดง กลางท่าไค (2531) ได้ประยุกต์ทฤษฎีการสรุปอ้างอิงในการหาความเชื่อมั่นของการประเมินความเที่ยงตรงเชิงเนื้อหา รูปแบบในการศึกษานั้น ผู้เชี่ยวชาญทุกคนประเมินความเที่ยงตรงของข้อสอบทุกข้อว่าวัดตรงจุดประสงค์เพียงใด รูปแบบการวัดในการศึกษา คือ $i \times (r : s)$ โดยผู้เชี่ยวชาญทุกคนประเมินข้อสอบทุกข้อ และผู้เชี่ยวชาญถูกสุ่มจากโรงเรียนแต่ละโรงเรียนจำนวนเท่า ๆ กัน เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบทดสอบวิชาคณิตศาสตร์ เรื่อง เซต และประเมินเชิงเนื้อหาชนิด 5 มาตรา พบว่า

1. ข้อสอบที่ผู้เชี่ยวชาญที่อยู่ในโรงเรียนและปฏิสัมพันธ์ระหว่างข้อสอบกับโรงเรียน เป็นแหล่งความแปรปรวนที่มีอิทธิพลต่อการประมาณค่าความเชื่อมั่นของการประเมินความเที่ยงตรงเชิงเนื้อหา โดยอาศัยผู้เชี่ยวชาญ

2. รูปแบบการวัด ที่ให้ค่าความเชื่อมั่นของการประเมินความเที่ยงตรงเชิงเนื้อหา โดย ผู้เชี่ยวชาญสูงสุดสำหรับรูปแบบความสัมพันธ์ระหว่างตัวประกอบ $i \times (r : s)$

3. สำหรับกลุ่มตัวอย่าง จะเป็นตัวแทนของสมาชิกทั้งหมดในเอกภพของตัวประกอบ แต่ละตัวที่พอจะหาได้ คือ ข้อสอบอย่างน้อย 9 ข้อ และผู้เชี่ยวชาญไม่เกิน 9 คน ต่อโรงเรียนโดยการสุ่มจากโรงเรียนอย่างน้อย 7 โรงเรียน จึงจะได้ค่าสัมประสิทธิ์การสรุปอ้างอิงอย่างน้อยมีค่าเท่ากับ .80

4. การเพิ่มขนาดของกลุ่มตัวอย่างของตัวประกอบแต่ละตัว ความแปรปรวนของความคลาดเคลื่อนมีค่าลดลงและสัมประสิทธิ์การสรุปอ้างอิงมีค่าเพิ่มขึ้นโดยเฉพาะการเพิ่มขนาดของกลุ่มตัวอย่างของตัวประกอบที่ต้องการสรุปอ้างอิง จะทำให้ค่าสัมประสิทธิ์การสรุปอ้างอิงมีค่าเพิ่มขึ้นมากกว่าการเพิ่มขนาดตัวอย่างของตัวประกอบอื่น และการสรุปอ้างอิงผลการวัดไปยังเอกภพจำกัด จะทำให้ค่าสัมประสิทธิ์การสรุปอ้างอิงสูงกว่าการสรุปอ้างอิงผลการวัดไปยังเอกภพไม่จำกัด

ไพรัตน์ วงษ์นาม (2533 : 105-123) ได้ศึกษาเปรียบเทียบผลของวิธีตรวจ การชี้แจงและการรู้ผลการเรียนของผู้ตอบ ที่มีต่อสัมประสิทธิ์การอ้างอิงสรุป (Generalizability Coefficient) ของการสอบแบบสอบความเรียงที่วัดความสามารถในการแสดงความคิดเห็นเกี่ยวกับข่าวและเหตุการณ์ ของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 30 คน ข้อสอบความเรียง 5 ข้อ และครูผู้ตรวจให้คะแนน จำนวน 20 คน สุ่มแบ่งผู้ตรวจให้คะแนนเป็น 2 กลุ่ม ๆ ละ 10 คน สำหรับการตรวจ 2 วิธี คือวิธีประเมินรวม และวิธีวิเคราะห์ ในแต่ละวิธีทำการตรวจ 3 ครั้ง คือ ครั้งแรกตรวจโดยใช้ประสบการณ์เดิมของผู้ตรวจให้คะแนน ครั้งที่สองตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยโดยไม่รู้ผลการเรียนของผู้ตอบ ครั้งที่สาม ตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยเหมือนครั้งที่สอง แต่รู้ผลการเรียนของผู้ตอบด้วย พบว่า เมื่ออ้างอิงสรุปไปยังเอกภพของข้อสอบและผู้ตรวจให้คะแนนพร้อมกัน ค่าสัมประสิทธิ์การอ้างอิงสรุป (ความเที่ยง) ของแบบสอบความเรียงที่ตรวจโดยวิธีประเมินรวม มีค่าอยู่ระหว่าง 0.3328 – 0.4782 และวิธีวิเคราะห์ มีค่าอยู่ระหว่าง 0.3348 – 0.5895 และใช้ข้อสอบ 6 ข้อ ผู้ตรวจให้คะแนนอย่างน้อย ในการตรวจวิธีวิเคราะห์ ตามที่ได้รับการชี้แจงและไม่รู้ผลการเรียนของผู้ตอบ จึงจะให้ค่าสัมประสิทธิ์การอ้างอิงสรุปไม่น้อยกว่า 0.50 เมื่ออ้างอิงสรุปไปยังเอกภพของผู้ตรวจให้คะแนนอย่างเดียว ค่าสัมประสิทธิ์การอ้างอิงสรุป (ความเที่ยง) ที่ตรวจโดยวิธีประเมินรวมมีค่าอยู่ระหว่าง 0.4743 – 0.6865 และวิธีวิเคราะห์มีค่าอยู่ระหว่าง 0.5985 – 0.7761 และใช้ข้อสอบ 5 ข้อ ผู้ตรวจให้คะแนน 5 คน ในการตรวจวิธีวิเคราะห์ตามที่ได้รับการชี้แจง และไม่รู้ผลการเรียนของผู้ตอบ จึงจะให้ค่าสัมประสิทธิ์การอ้างอิงสรุปไม่น้อยกว่า 0.50 เมื่ออ้างอิงไปยังเอกภพของผู้ตรวจให้คะแนนอย่างเดียว พบว่า เฉพาะกรณีที่ผู้ตรวจให้คะแนนได้รับการชี้แจงและไม่รู้ผลการเรียนของผู้ตรวจให้คะแนนเท่านั้นให้ค่าสัมประสิทธิ์การอ้างอิงสรุปสูงกว่าการตรวจโดยใช้ประสบการณ์เดิม ถ้ารู้ผลการเรียนของผู้ตอบค่าสัมประสิทธิ์การอ้างอิงสรุปจะไม่แตกต่างกัน

กนกวรรณ เอี่ยมชัย (2539 : บทคัดย่อ) ศึกษาความเชื่อมั่นของแบบทดสอบเอ็มอีคิววิชาแนวคิดพื้นฐานและหลักการพยาบาล ที่ได้จากวิธีการตรวจให้คะแนนต่างกัน กลุ่มตัวอย่างเป็นนักศึกษาพยาบาล ชั้นปีที่ 2 ของวิทยาลัยการพยาบาล สังกัดสถาบันราชชนก จำนวน 250 คน เครื่องมือที่ใช้เป็นแบบทดสอบเอ็มอีคิววิชาแนวคิดพื้นฐานและหลักการพยาบาลที่ผู้วิจัยได้สร้างขึ้น จำนวน 4 กรณีศึกษา ข้อคำถามทั้งหมด 20 ข้อ คำนวณค่าสัมประสิทธิ์การสรุปอ้างอิงตามวิธีของเบรนนอนและทดสอบความแตกต่างของสัมประสิทธิ์การสรุปอ้างอิงของวิธีการตรวจให้คะแนนทั้งสองวิธีด้วยสูตร UX, ของวูดรฟฟ์และเฟลด์ทดสอบความเชื่อมั่นของที่ได้จากการตรวจให้คะแนนทั้งสองวิธีว่าอยู่ในช่วงค่าพารามิเตอร์ที่แท้จริง ด้วยสูตร t ของ คริสตอฟ พบว่า

1. ความแปรปรวนผลรวมระหว่างข้อสอบและผู้ตรวจ ที่ได้จากวิธีการตรวจให้คะแนนทั้งสองวิธี มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบมากที่สุด

2. ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบที่ได้จากวิธีการตรวจของน็อก มีค่าสูงกว่าวิธีการตรวจของเฟลเดติ และแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01

3. ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบที่ได้จากการตรวจให้คะแนนทั้งสองวิธีอยู่ในช่วงค่าพารามิเตอร์ที่แท้จริงทั้งสองค่า

4. เมื่อกำหนดให้ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบมีค่า .80 เป็นอย่างต่ำ วิธีการตรวจให้คะแนนของน็อกใช้ข้อสอบจำนวน 10 ข้อ ต้องใช้ผู้ตรวจจำนวน 4 คน ใช้ข้อสอบจำนวน 15 ข้อ ต้องใช้ผู้ตรวจจำนวน 3 คน ใช้ข้อสอบจำนวน 20-30 ข้อ ต้องใช้ผู้ตรวจจำนวน 12 คน ใช้ข้อสอบจำนวน 35 ข้อ ต้องใช้ผู้ตรวจจำนวน 1 คน และวิธีการตรวจของ เฟลเดติ ใช้ข้อสอบจำนวน 15 ข้อ ต้องใช้ผู้ตรวจจำนวน 6 คน ใช้ข้อสอบจำนวน 20 ข้อ ต้องใช้ผู้ตรวจจำนวน 4 คน ใช้ข้อสอบจำนวน 25-30 ข้อ ต้องใช้ผู้ตรวจจำนวน 3 คน ใช้ข้อสอบจำนวน 35 ข้อ ต้องใช้ผู้ตรวจจำนวน 2 คน และถ้าใช้ผู้ตรวจจำนวน 1 คน ไม่สามารถนำไปปฏิบัติได้

5. เมื่อใช้ข้อสอบจำนวน 5 ข้อ ต้องใช้ผู้ตรวจจำนวน 1-2 คน และใช้ข้อสอบจำนวน 10 ข้อ ต้องใช้ผู้ตรวจจำนวน 1 คน ค่าความเชื่อมั่นของแบบทดสอบที่ได้จากการตรวจให้คะแนนทั้งสองวิธี แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และถ้าใช้ข้อสอบจำนวน 5 ข้อ ต้องใช้ผู้ตรวจจำนวน 3-6 คน ใช้ข้อสอบจำนวน 10 ข้อ ต้องใช้ผู้ตรวจจำนวน 2-6 คน ข้อสอบจำนวน 15-35 ข้อ ต้องใช้ผู้ตรวจจำนวน 1-6 คน ความเชื่อมั่นของแบบทดสอบที่ได้จากการตรวจให้คะแนนทั้งสองวิธีแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติที่ระดับ .01

ชาญ วชิรเดช (2538) ศึกษารูปแบบของแบบทดสอบการแก้ปัญหาที่สอบด้วยไมโครคอมพิวเตอร์ 3 รูปแบบ คือ แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการเพียงอย่างเดียว ประเมิน 4 ตัวเลือก แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการและข้อเท็จจริง ประเมิน 4 ตัวเลือก แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการเพียงอย่างเดียว ประเมินเลือกตอบหลายตัวเลือก แบบทดสอบแต่ละรูปแบบประกอบด้วยเงื่อนไขสถานการณ์ของโจทย์ปัญหา 6 เงื่อนไข แต่ละเงื่อนไขประกอบด้วยคำถามย่อย 5 ข้อ เป็นข้อสอบทั้งสิ้น 30 ข้อ พบว่า ความแปรปรวนของผลรวมระหว่างนักเรียนทุกคนทำข้อสอบทุกข้อที่แบ่งอยู่ในเงื่อนไขสถานการณ์ของโจทย์ที่กำหนด มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบการแก้ปัญหาที่สอบด้วยไมโครคอมพิวเตอร์ทั้งสามรูปแบบ สัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบทั้งสามรูปแบบแตกต่างกันอย่างมีนัยสำคัญทางสถิติ ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบที่กำหนดเงื่อนไขสถานการณ์ของโจทย์ปัญหาที่กำหนด ตั้งแต่ 1 เงื่อนไขโดยใช้จำนวนข้อสอบ 1-10 ข้อ แบบทดสอบความสามารถในการแก้ปัญหาที่อาศัยหลักการและข้อเท็จจริงชนิดประเมิน 4 ตัวเลือก จะมีค่าสัมประสิทธิ์การสรุปอ้างอิงสูงกว่ารูปแบบอื่นๆ เงื่อนไข และค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบทั้งสามรูปแบบมีค่าสัมประสิทธิ์การสรุปอ้างอิงประมาณใกล้เคียงกับค่าพารามิเตอร์ของความเชื่อมั่นที่แท้จริง

✓ รัชกร สุวรรณจรัส (2540 : บทความย่อ) ศึกษาความเชื่อมั่นของแบบทดสอบเอ็มอีคิววัดความสามารถในการแก้ปัญหาวิชาบัญชีกิจการบริการที่มีผู้ตรวจให้คะแนน 1 คน กลุ่ม ตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 ของกลุ่มโรงเรียนกรมสามัญศึกษาส่วนกลางกลุ่มที่ 5 จำนวน 180 คน เครื่องมือที่ใช้เป็นแบบทดสอบเอ็มอีคิววัดความสามารถในการแก้ปัญหาวิชาบัญชีกิจการบริการจำนวน 2 ฉบับ ที่ผู้วิจัยสร้างขึ้นคำนวณค่าสัมประสิทธิ์การสรุปอ้างอิงและการตัดสินใจสรุปอ้างอิงตามวิธีของเบรนนอนด้วยโปรแกรมสำเร็จรูป GENOVA และทดสอบความแตกต่างของค่าความเชื่อมั่นด้วยสูตร U_X , ผลการวิจัยพบว่า

1. ค่าความเชื่อมั่นของแบบทดสอบทั้งสองฉบับที่คำนวณด้วยสูตร r_{F-R} ของ Feldt-Raju แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

2. ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบทั้งสองฉบับ แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

3. ในการตัดสินใจการสรุปอ้างอิงเมื่อกำหนดให้ค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบมีค่า .70 เป็นอย่างต่ำ มีผู้ตรวจให้คะแนน 1 คน แบบทดสอบฉบับ A ต้องใช้ ข้อสอบ 16 ข้อขึ้นไป และฉบับ B ต้องใช้ข้อสอบ 10 ข้อขึ้นไป

4. ในการคำนวณค่าความเชื่อมั่นของแบบทดสอบฉบับเดียวกันด้วยสูตร r_{F-R} ของ Feldt-Raju กับสูตรสัมประสิทธิ์การสรุปอ้างอิงของเบรนนอน สูตร r_{F-R} จะให้ค่าความเชื่อมั่นสูงกว่า และเมื่อทดสอบความแตกต่างทั้งสองสูตรมีค่าไม่แตกต่างกันทางสถิติ

อุษณีย์ บัวศิริพันธุ์ (2543) ศึกษาเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจให้คะแนน และประสิทธิภาพของผู้ตรวจให้คะแนน แตกต่างกัน กรณีผลการศึกษาเกี่ยวกับวิธีการตรวจให้คะแนน การตรวจแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนนของผู้ตรวจให้คะแนนเอง เมื่อมีจำนวนผู้ตรวจให้คะแนน 3 , 4 , 5 และ 6 คน มีคะแนนเฉลี่ยอยู่ระหว่าง 8.79 – 9.05 โดยการตรวจให้คะแนนเมื่อมีจำนวนผู้ตรวจให้คะแนน 4 คน คะแนนเฉลี่ยมีแนวโน้มสูงที่สุด รองลงมาคือ มีจำนวนผู้ตรวจให้คะแนน 5 คน และเมื่อมีจำนวนผู้ตรวจให้คะแนน 6 คน คะแนนเฉลี่ยมีแนวโน้มต่ำที่สุด ส่วนการตรวจให้คะแนนแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนนแบบรูปิก เมื่อมีจำนวนผู้ตรวจให้คะแนน 3 , 4 , 5 และ 6 มีคะแนนเฉลี่ยอยู่ระหว่าง 9.82 – 10.05 โดยการตรวจให้คะแนนเมื่อมีจำนวนผู้ตรวจให้คะแนน 4 คน คะแนนเฉลี่ยมีแนวโน้มสูงที่สุด รองลงมาคือ จำนวนผู้ตรวจให้คะแนน 3 และ 6 คน และเมื่อมีจำนวนผู้ตรวจให้คะแนน 5 คน คะแนนเฉลี่ยมีแนวโน้มต่ำที่สุด

จากการศึกษาผลงานวิจัยทั้งในด้าน ทักษะกระบวนการทางวิทยาศาสตร์ การสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ ความเชื่อมั่นของการประเมินทักษะกระบวนการทางวิทยาศาสตร์ ความเชื่อมั่นของการประเมินผลในรายวิชาอื่น ๆ ที่เกี่ยวข้องกัน รวมทั้งความเชื่อมั่นของการประเมินผลตามทฤษฎีการสรุปอ้างอิง (Generalizability Theory) สรุปได้ว่า การพัฒนาทฤษฎีการให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ยังไม่เป็นที่แพร่หลาย ยังไม่สามารถนำไปใช้ในการปฏิบัติจริงภายในห้องเรียน โดยเฉพาะค่าความเชื่อมั่นของกฎเกณฑ์การให้คะแนน (Scoring Rubric) ครูผู้สอนยังขาดความมั่นใจ เมื่อนำไปใช้ในห้องเรียน ผู้วิจัยจึงได้เลือกการพัฒนาทฤษฎีการให้คะแนน (Scoring Rubric) เป็นหัวข้อในการวิจัย และศึกษารูปแบบของการตรวจให้คะแนนโดยวิเคราะห์จากค่าสัมประสิทธิ์การสรุปอ้างอิง

สมมุติฐานการวิจัย

ค่าความเชื่อมั่นของการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ใช้รูปแบบการตรวจต่างกันให้ค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการวิจัย

การกำหนดประชากรและเลือกกลุ่มตัวอย่าง

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนมัธยมศึกษาสังกัดกรมสามัญศึกษาจังหวัดนครราชสีมา ปีการศึกษา 2544 จำนวน 16,919 คน

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนบุญวัฒนา 2 สังกัดกรมสามัญศึกษาจังหวัดนครราชสีมา ปีการศึกษา 2544 จำนวน 50 คน ซึ่งเลือกมาโดยการสุ่มแบบหลายขั้นตอน (Multi - Stage Random Sampling) และเนื่องจาก ผู้วิจัยมีจุดมุ่งหมายที่จะสร้างและพัฒนา กฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ และศึกษาผลจากการวิเคราะห์ค่าความเชื่อมั่นของผู้ตรวจให้คะแนนที่อ้างอิงผลการตรวจไปสู่ค่าความเชื่อมั่นของกฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ ดังนั้น การกำหนดขนาดของกลุ่มตัวอย่างจึงต้องคำนึงถึงความเหมาะสม ซึ่งผู้ตรวจสามารถนำไปใช้ประมาณค่าสัมประสิทธิ์การสุบอ้างอิงได้อย่างมั่นใจ สมิท (Smith. 1978 : 319-346) ได้เสนอขนาดของกลุ่มตัวอย่างขั้นต่ำที่ใช้ศึกษาทั้ง 3 องค์ประกอบ คือ $n_p \times n_i \times n_r$ ควรมีอย่างน้อย 1,000 ค่า ดังนั้น ผู้วิจัยจึงขอกำหนดจำนวนผู้เข้าสอบ (n_p) 50 คน จำนวนข้อสอบ (n_i) 5 ทักษะและจำนวนผู้ตรวจ (n_r) 5 คน ได้ขนาดของกลุ่มตัวอย่าง $50 \times 5 \times 5 = 1,250$ ค่า เป็นไปตามกฎเกณฑ์ของสมิท ขั้นตอนในการสุ่มมีดังนี้

ขั้นที่ 1 ประมาณโรงเรียนที่เป็นกลุ่มตัวอย่างที่ใช้ในการวิจัย ด้วยการเลือกสุ่มแบบเจาะจง (Purposive Sampling) เลือกโรงเรียนบุญวัฒนา 2 เป็นกลุ่มตัวอย่างที่ใช้ในการวิจัย เนื่องจากผลการทดสอบของนักเรียนเป็นเพียงข้อมูลประกอบในการวิจัยเท่านั้น ข้อมูลหลักที่ใช้ในการวิจัยคือ ผลการตรวจให้คะแนนของผู้ตรวจที่จะนำมาพิจารณาความสอดคล้อง และประมาณค่าสัมประสิทธิ์การสุบอ้างอิง ผู้วิจัยจึงตัดสินใจเลือกกลุ่มตัวอย่างเพียงโรงเรียนเดียวเพื่อใช้ในการวิจัย

ขั้นที่ 2 ใช้วิธีการสุ่มอย่างง่าย (Simple Random Sampling) แบ่งกลุ่มตัวอย่างออกเป็น 2 กลุ่ม โดยใช้วิธีการจับฉลาก คือ

1. กลุ่มตัวอย่างที่ใช้ในการสร้างและหาคุณภาพของเครื่องมือ จำนวน 20 คน
2. กลุ่มตัวอย่างที่ใช้ในการหาความเชื่อมั่นของกฎเกณฑ์การให้คะแนนและศึกษารูปแบบของการตรวจโดยใช้ทฤษฎีการสุบอ้างอิง จำนวน 50 คน

รูปแบบการทดลองที่ใช้ในการวิจัย

การศึกษาค้นคว้าในครั้งนี้ ผู้วิจัยกำหนดความสัมพันธ์ขององค์ประกอบที่ศึกษา ในชั้นการศึกษาเพื่อ การสรุปอ้างอิงเป็น 2 รูปแบบ คือ

1. แบบแผน $i \times r \times p$ เป็นแบบแผนการวิเคราะห์ค่าความเชื่อมั่นของคะแนนแบบทดสอบวัดทักษะ กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ที่ได้จาก ผู้เข้าสอบทุกคนสอบด้วยแบบทดสอบชุดเดียวกัน และ ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบทุกคน

2. แบบแผน $i \times (p : r)$ เป็นแบบแผนการวิเคราะห์ค่าความเชื่อมั่นของคะแนนแบบทดสอบวัดทักษะ กระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ที่ได้จาก ผู้เข้าสอบทุกคนสอบด้วยแบบทดสอบชุดเดียวกัน และ ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบบางคน แต่แยกตรวจผู้เข้าสอบเป็นรายกลุ่มโดยไม่ซ้ำ กัน

เครื่องมือที่ใช้ในการวิจัย

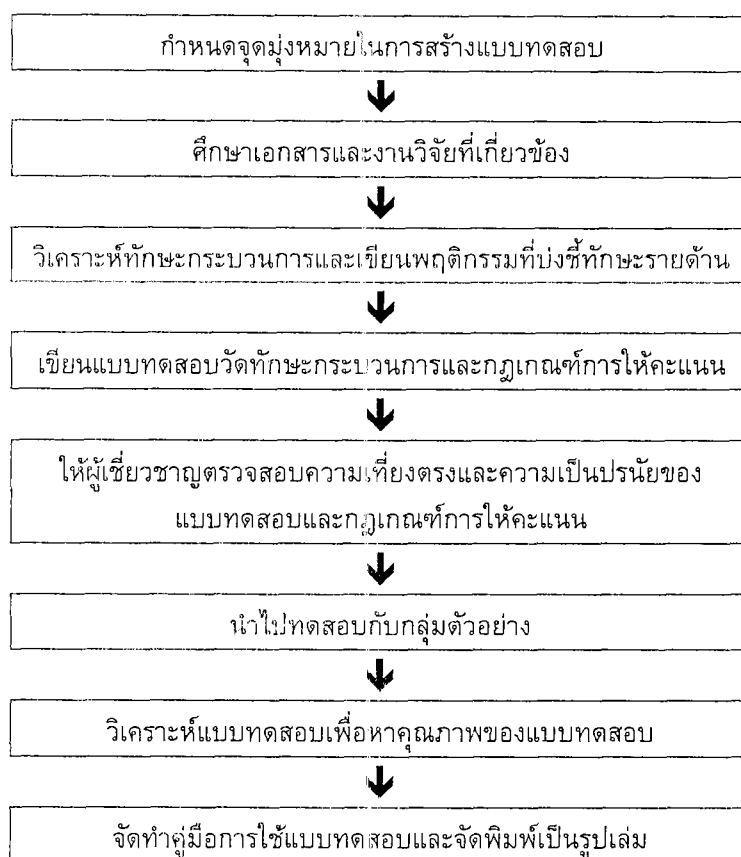
เครื่องมือที่ใช้ในการศึกษาค้นคว้าในครั้งนี้ เป็นแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ ขั้นบูรณาการ มีลักษณะเป็นคำถามปลายเปิดและกำหนดสถานการณ์ที่เกี่ยวกับวิทยาศาสตร์ให้ตอบประกอบ ด้วย 5 ทักษะ คือ

- | | |
|---|-------------|
| 1. ทักษะการตั้งสมมุติฐาน | จำนวน 3 ข้อ |
| 2. ทักษะการกำหนดนิยามเชิงปฏิบัติการ | จำนวน 3 ข้อ |
| 3. ทักษะการกำหนดและควบคุมตัวแปร | จำนวน 4 ข้อ |
| 4. ทักษะการทดลอง | จำนวน 1 ข้อ |
| 5. ทักษะการตีความหมายข้อมูลและลงข้อสรุป | จำนวน 4 ข้อ |

รวมเป็นแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 1 ฉบับ จำนวน 5 ทักษะ จำนวนข้อสอบรวม 15 ข้อ ใช้เวลา 120 นาที

การสร้างเครื่องมือที่ใช้ในการวิจัย

ในการสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้



ภาพประกอบ 3 แสดงลำดับในการสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์

1. กำหนดจุดมุ่งหมายในการสร้างแบบทดสอบ

1.1 เพื่อสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 5 ทักษะลักษณะเป็นคำถามปลายเปิดและกำหนดสถานการณ์ให้ตอบ สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย

1.2 เพื่อพัฒนากฎเกณฑ์การให้คะแนน (Scoring Rubric) ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

2. ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับความหมาย ประเภทของทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการและการประเมินผลทักษะกระบวนการทางวิทยาศาสตร์

2.2 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับความหมาย วิธีการสร้าง และศึกษาอิทธิพลที่มีต่อการวัดผลโดยใช้แบบทดสอบความเรียง

2.3 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับความหมาย ประเภท วิธีการสร้างและการตรวจให้คะแนนโดยใช้กฎเกณฑ์การให้คะแนน (Scoring Rubric)

2.4 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับทฤษฎีค่าความเชื่อมั่น ได้แก่ การประมาณค่าความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม การประมาณค่าความเชื่อมั่นที่เกี่ยวกับการตรวจให้คะแนน และการประมาณค่าความเชื่อมั่นตามโครงสร้างทฤษฎีการสรุปอ้างอิง เพื่อใช้เป็นแนวทางในการดำเนินการ

3. วิเคราะห์ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการและเขียนพฤติกรรมที่บ่งชี้ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการย่อยทั้ง 5 ทักษะ กำหนดการวัดในแต่ละทักษะโดยเชิญครูผู้สอนวิทยาศาสตร์ซึ่งถือว่าเป็นผู้เชี่ยวชาญมาร่วมกันวิเคราะห์ ผลการวิเคราะห์ดังแสดงในตาราง 11

ตาราง 11 ผลการวิเคราะห์พฤติกรรมที่บ่งชี้ทักษะกระบวนการทางวิทยาศาสตร์

ทักษะ	พฤติกรรมที่บ่งชี้
- การตั้งสมมุติฐาน	คาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตามโดยอาศัยความรู้และประสบการณ์เดิม
- การกำหนดนิยามเชิงปฏิบัติการ	การกำหนดความหมายและขอบเขตของคำหรือตัวแปรต่าง ๆ ในปฏิบัติการทดลองให้สามารถทดสอบหรือวัดได้
- การกำหนดและควบคุมตัวแปร	การบ่งชี้และกำหนดตัวแปรต้น ตัวแปรตาม และควบคุมตัวแปรที่เป็นสาเหตุอื่น ๆ ในปรากฏการณ์หนึ่ง ๆ ที่ต้องการศึกษา
- การทดลอง	การออกแบบการทดลอง ระบุวัสดุอุปกรณ์ที่ใช้ในการทดลอง อธิบายขั้นตอนการทดลอง และเสนอแนะการบันทึกผลการทดลอง
- การตีความหมายข้อมูลและลงข้อสรุป	การแปลความหมายหรือบรรยายลักษณะและคุณสมบัติของข้อมูล พร้อมทั้งสรุปความสัมพันธ์ของตัวแปรที่อยู่ในข้อมูลที่ต้องการศึกษา

4. เขียนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่เป็นแบบความเรียง ตามพฤติกรรมที่บ่งชี้ในตาราง 11 จำนวน 19 ข้อ เพื่อทดสอบและหาคุณภาพรายข้อ และคัดเลือกไว้ 15 ข้อ การกำหนดน้ำหนักคะแนนในแต่ละข้อให้เป็นไปตามข้อย่อย พร้อมทั้งจัดทำคู่มือการตรวจให้คะแนน (เฉลย) และสร้างกฎเกณฑ์การให้คะแนนแบบเฉพาะ (Specific Rubric) แยกเป็นรายทักษะ เพื่อให้ผู้ตรวจใช้ในการให้คะแนนแต่ละทักษะทั้งสองรูปแบบ

5. นำข้อคำถามที่ผู้วิจัยสร้างขึ้นไปให้ผู้เชี่ยวชาญ ซึ่งประกอบด้วยผู้เชี่ยวชาญทางด้านการวัดผลที่มีวุฒิตั้งแต่ปริญญาโทขึ้นไป จำนวน 4 ท่าน และผู้เชี่ยวชาญด้านการสอนวิทยาศาสตร์ จำนวน 1 ท่าน ที่มีวุฒิตั้งแต่ปริญญาโทขึ้นไป ดังรายชื่อในภาคผนวก ตรวจสอบความเที่ยงตรงเชิงเนื้อหา (Content Validity) และ

ความเป็นปรนัย (Objectivity) ของข้อคำถามและคู่มือการตรวจให้คะแนน นำผลการวิเคราะห์ความเที่ยงตรงเชิงเนื้อหาโดยใช้ดัชนีความสอดคล้อง (IOC) โดยรวมความเห็นในช่อง เห็นด้วย (+1) ไม่แน่ใจ (0) และไม่เห็นด้วย (-1) ตามวิธีของโรวินสกีและแฮมเบลตัน (บุญเชิด ภิญโญนนตพงษ์. 2526) คัดเลือกข้อที่มีค่า IOC ตั้งแต่ 0.5 ขึ้นไป ปรากฏผลการคำนวณค่า IOC มีค่าระหว่าง 0.6 – 1.0 ดังแสดงในภาคผนวก

6. นำแบบทดสอบที่ผ่านการตรวจสอบคุณภาพในขั้นต้นแล้ว ไปทดสอบกับกลุ่มตัวอย่างที่ใช้ในการหาคุณภาพของเครื่องมือ คือ นักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนบุญวัฒนา 2 จำนวน 20 คน โดยการให้นักเรียนทำข้อสอบที่ละข้อเพื่อตรวจสอบเวลาที่เหมาะสมในการทำข้อสอบ โดยพิจารณาจากจำนวนนักเรียนที่ทำได้เร็วหรือช้าประมาณร้อยละ 80 ของนักเรียนที่เข้าสอบ

7. นำผลการทดสอบมาตรวจให้คะแนนโดยผู้วิจัยเป็นผู้ดำเนินการเอง นำผลการตรวจให้คะแนนมาวิเคราะห์เพื่อคัดเลือกข้อสอบและปรับปรุง ได้ค่าความยากอยู่ระหว่าง .246 - .789 และค่าอำนาจจำแนกเป็นรายข้อ มีค่าอยู่ระหว่าง .332 – .845 โดยใช้ r_{xy} ได้ข้อสอบจำนวน 15 ข้อ คือ ข้อ 1.1, ข้อ 1.2, ข้อ 1.3, ข้อ 2.1, ข้อ 2.2, ข้อ 2.4, ข้อ 3.1, ข้อ 3.2, ข้อ 3.3, ข้อ 3.4, ข้อ 4.3, ข้อ 5.1, ข้อ 5.2, ข้อ 5.3 และ ข้อ 5.4 ซึ่งข้อที่ผ่านการคัดเลือกมีค่าความยากอยู่ระหว่าง .316 - .789 และค่าอำนาจจำแนกมีค่าอยู่ระหว่าง .529 – .845 ดังแสดงในตารางภาคผนวก

8. จัดทำคู่มือดำเนินการสอบและจัดพิมพ์

ตัวอย่างเครื่องมือที่ใช้ในการวิจัย

แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

ข้อ 1 ทักษะการตั้งสมมุติฐาน

1.1 จากการทดลองละลายสาร A ในน้ำเย็นปรากฏว่า สาร A 21 กรัม ละลายได้พอดีในน้ำ 250 ลูกบาศก์เซนติเมตร จากการทดลองนี้ ถ้าผู้ทดลองมีความเชื่อว่า การอุ่นสารละลายให้ร้อนขึ้นจะช่วยให้สาร A ละลายได้เพิ่มขึ้นอีก ผู้ทดลองควรตั้งสมมุติฐานว่าอย่างไร

.....

.....

.....

1.2 นำลูกกลมตันที่ทำด้วยวัตถุต่างชนิดกัน ได้แก่ วัตถุ A , B และวัตถุ C ซึ่งมีเส้นผ่าศูนย์กลางเท่ากันมา 3 ลูก ปล่อยให้กลิ้งบนพื้นเอียงเดียวกัน แล้วจับเวลาที่ลูกกลมแต่ละชนิดใช้ในการกลิ้งจากปลายบนสุดถึงปลายล่างสุดของพื้นเอียง ทำซ้ำแบบเดิม 3 ครั้ง หาค่าเฉลี่ยและนำมาเปรียบเทียบกัน กรณีดังกล่าว ถ้านักเรียนเป็นผู้ทดลองจะตั้งสมมุติฐานอย่างไร

.....

.....

.....

1.3 จงศึกษาข้อมูลจากตารางต่อไปนี้

ชนิดของสาร	ความสามารถในการละลายของสารที่อุณหภูมิห้อง (กรัม)	
	ในน้ำ 100 ลบ.ซม.	ในเอทิลแอลกอฮอล์ 100 ลบ.ซม.
A	140	5
B	39	1
C	200	0.6
D	36	0
E	6	0

ข้อมูลในการทดลองจากตารางนี้ ควรตั้งสมมุติฐานอย่างไร

.....

.....

.....

แนวทางในการตอบ

ข้อ 1 ทักษะการตั้งสมมุติฐาน

- 1.1 การเพิ่มอุณหภูมิให้แก่ไอจะมีผลต่อปริมาณการละลายของสาร A
- 1.2 ชนิดของวัตถุที่ใช้ทำลูกกลมนั้นมีผลต่อเวลาในการกลิ้งของลูกกลมนบนพื้นเอียง
- 1.3 ความสามารถในการละลายของสารแต่ละชนิดขึ้นอยู่กับชนิดของตัวทำละลาย หรือ ชนิดของตัว

ทำละลายมีผลต่อความสามารถในการละลายของสารแต่ละชนิด

กฎเกณฑ์การให้คะแนนทักษะการตั้งสมมุติฐาน

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดชัดเจนถูกต้องทั้งหมด
3 ดี	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดชัดเจนถูกต้อง 2 ข้อใน 3 ข้อ
2 ใช้ได้	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดชัดเจนถูกต้อง 1 ข้อใน 3 ข้อ
1 ใช้ไม่ได้	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองไม่สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดมีแนวโน้มที่จะได้คำตอบ
0	ไม่ตอบหรือตอบไม่ถูกเลย

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลในการวิจัย ดังนี้

1. นำหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์กับโรงเรียนที่เป็นกลุ่มตัวอย่าง
2. ติดต่อประสานงาน ขออนุญาตผู้บริหารและครูผู้สอนของโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลา สถานที่ วิธีดำเนินการสอบ และจัดประชุมชี้แจงครูที่ได้รับการคัดเลือกให้มาตรวจให้คะแนนจำนวน 5 คน ในครั้งนี้ เพื่อให้เกิดความเข้าใจอันดี
3. นำแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผ่านการวิเคราะห์คุณภาพรายข้อแล้ว ไปทดสอบกับกลุ่มตัวอย่าง จำนวน 50 คน ผู้ดำเนินการสอบ 2 คน ซึ่งผู้ดำเนินการสอบชี้แจงการปฏิบัติในการดำเนินการสอบ ดังนี้
 - 3.1 อธิบายให้ผู้เข้าสอบกลุ่มตัวอย่างเข้าใจวัตถุประสงค์ และผลประโยชน์ที่จะได้รับการทำแบบทดสอบ
 - 3.2 อธิบายให้ผู้เข้าสอบทุกคนเข้าใจถึงวิธีดำเนินการสอบ วิธีการปฏิบัติ และวิธีตอบ ก่อนที่จะให้ผู้เข้าสอบลงมือทำ

การจัดกระทำและการวิเคราะห์ข้อมูล

1. นำกระดาษคำตอบของผู้เข้าสอบทั้งหมดมาถ่ายเอกสารจำนวน 6 ชุด จัดให้ผู้ตรวจจำนวน 5 คน ได้ตรวจให้คะแนนตามแนวทางในการตอบและกฎเกณฑ์การให้คะแนนที่ผู้วิจัยได้สร้างขึ้น ในคู่มือการตรวจให้คะแนน โดยแบ่งการดำเนินการเป็น 2 วิธีดังนี้
 - 1.1 นำกระดาษคำตอบชุดที่ 1 มาจัดแบ่งออกเป็น 5 กลุ่ม ๆ ละ 10 คน นำไปให้ผู้ตรวจให้คะแนนแยกกันตรวจ 1 คน ต่อ 1 กลุ่ม (ตามรูปแบบผู้ตรวจทุกคนตรวจให้คะแนนแบบทดสอบทุกทักษะของผู้เข้าสอบบางคน) แล้วบันทึกผลการตรวจให้คะแนน
 - 1.2 หลังจากนั้น 1 สัปดาห์ นำกระดาษคำตอบจำนวน 5 ชุด ที่เหลือไปให้ผู้ตรวจทุกคนได้ตรวจให้คะแนนคนละ 1 ชุด (ตามรูปแบบผู้ตรวจทุกคนตรวจให้คะแนนแบบทดสอบทุกทักษะของผู้เข้าสอบทุกคน) แล้วบันทึกผลการตรวจให้คะแนน
2. นำผลที่ได้จากการตรวจให้คะแนนมาวิเคราะห์ข้อมูล ดังนี้
 - 2.1 คำนวณค่าสถิติพื้นฐานของแบบทดสอบจำแนกตามผู้ตรวจให้คะแนน ตามข้อ 1.2
 - 2.2 คำนวณค่าความเชื่อมั่นของแบบทดสอบ โดยใช้สูตร r_{F-R} ของ Feldt - Raju (บุญเชิด ภิญโญนนตพงษ์. 2545)
 - 2.3 คำนวณหาความสอดคล้องของการตรวจให้คะแนน โดยใช้สัมประสิทธิ์การสรุปอ้างอิงจากสูตรของเบรนนอน วิเคราะห์โดยใช้โปรแกรมสำเร็จรูป GENOVA มีรูปแบบในการคำนวณ ดังนี้
 - 2.3.1 นำผลการตรวจให้คะแนนในข้อ 1.2 มาคำนวณสัมประสิทธิ์การสรุปอ้างอิงตามแบบแผน $i \times r \times p$ ซึ่งจัดเป็นรูปแบบ Two – Facet Crossed Design

2.3.2 นำผลการตรวจให้คะแนนในข้อ 1.1 มาคำนวณสัมประสิทธิ์การสรุปอ้างอิงตามแบบแผน $I \times (p : r)$ ซึ่งจัดเป็นรูปแบบ Two – Facet Nested Design

2.4. ทดสอบความแตกต่างของผลการตรวจให้คะแนนทั้งสองรูปแบบ (ข้อ 2.3.1 และ 2.3.2) ตามสมมุติฐานของการวิจัย โดยใช้สูตร UX_1 ของวูดรUFF และเฟลด์ (Woodruff and Feldt. 1986 : 393-413)

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของแบบทดสอบ ได้แก่ ค่าเฉลี่ย ค่าส่วนเบี่ยงเบนมาตรฐาน และค่าสัมประสิทธิ์ความแปรผัน (Coefficient of Variation)

2. ค่าอำนาจจำแนกของข้อสอบรายข้อ โดยใช้สูตรสหสัมพันธ์แบบเพียร์สัน (Pearson Product Moment Correlation : r_{xy}) (บุญเชิด ภิญญอนันตพงษ์. 2545 : 84)

$$r_{xy} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

เมื่อ	r_{xy}	แทน	สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน หรือ ดัชนีอำนาจจำแนกของข้อสอบรายข้อ
	N	แทน	จำนวนกลุ่มตัวอย่าง
	$\sum X$	แทน	ผลรวมของคะแนนรายข้อ (item)
	$\sum Y$	แทน	ผลรวมของคะแนนรวม (total) ที่ไม่รวมข้อที่นำมาคำนวณ
	$\sum XY$	แทน	ผลรวมของผลคูณของคะแนน X และคะแนน Y
	$\sum X^2$	แทน	ผลรวมกำลังสองของคะแนน X
	$\sum Y^2$	แทน	ผลรวมกำลังสองของคะแนน Y

3. ค่าความเชื่อมั่น (Reliability) ของแบบทดสอบทั้งฉบับจากสูตร r_{F-R} ของ Feldt – Raju (บุญเชิด ภิญญอนันตพงษ์. 2545 : 139)

$$r_{F-R} = \frac{1}{[1 - \sum \lambda_i^2]} \cdot \frac{[S_X^2 - \sum S_i^2]}{S_X^2}$$

เมื่อ	r_{F-R}	แทน	ความเชื่อมั่นของแบบทดสอบ
	λ_i^2	แทน	สัดส่วนระหว่างผลรวมของเมตริกในแถวที่ i กับความแปรปรวนของแบบทดสอบทั้งฉบับ
	S_i^2	แทน	ความแปรปรวนของคะแนนในแต่ละด้าน
	S_X^2	แทน	ความแปรปรวนของคะแนนของแบบทดสอบทั้งฉบับ

4. ค่าความเชื่อมั่นที่ปรับขยายของการตรวจให้คะแนน วิเคราะห์ข้อมูลโดยใช้สูตรของสเปียร์แมน - บราวน์ (บุญเชิด ภิญโญนนตพงษ์. 2545 : 150)

$$r_{nn} = \frac{nr_{00}}{1 + (n-1)r_{00}}$$

เมื่อ	r_{nn}	แทน	ค่าความเชื่อมั่นที่ปรับขยายของการตรวจให้คะแนน
	n	แทน	จำนวนผู้ตรวจให้คะแนน
	r_{00}	แทน	ค่าความเชื่อมั่นเฉลี่ยของการตรวจให้คะแนน

5. ค่าความคลาดเคลื่อนมาตรฐานในการวัด (Standard error of measurement) ของแบบทดสอบ ทั้งฉบับ (บุญเชิด ภิญโญนนตพงษ์. 2545 : 117)

$$S_E = S_X \sqrt{1 - r_{tt}}$$

เมื่อ	S_E	แทน	ค่าความคลาดเคลื่อนมาตรฐานในการวัด
	S_X	แทน	ค่าความเบี่ยงเบนมาตรฐาน
	r_{tt}	แทน	ค่าความเชื่อมั่นของแบบทดสอบ

6. ค่าสัมประสิทธิ์สรูปอ้างอิง ($E\rho^2$) โดยใช้สูตรของเบรนนอน (Brennan. 1983 : 60)

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

เมื่อ	$E\rho^2$	แทน	ค่าสัมประสิทธิ์การสรูปอ้างอิง
	$\sigma^2(\rho)$	แทน	ความแปรปรวนของคะแนนเอกภพ
	$\sigma^2(\delta)$	แทน	ความแปรปรวนของคะแนนความคลาดเคลื่อน

7. สถิติทดสอบนัยสำคัญของความแตกต่างของค่าความเชื่อมั่น วิเคราะห์ข้อมูลโดยใช้สูตร UX_1 ของ วูดรUFF และ เฟลด์ (Woodruff and Feldt. 1986 : 393-413)

$$UX_1 = \frac{\sum_{i=1}^m (u_i - \bar{u})^2}{S_u^2 - C_u}$$

$$\bar{u} = \frac{\sum_{i=1}^m u_i}{m}$$

$$u_i = \frac{1}{(1 - r_i)^3}$$

$$S_u^2 = \frac{2}{9m(N_c - 1)} \sum_{i=1}^{m'} u_i^2$$

$$C_u = \frac{4}{9m(m-1)(N_c - 1)} \sum_{i=2}^m \sum_{j=1}^{i-1} [r(x_i, x_j)]^2 u_i u_j$$

$$N_c = N \left(\frac{n-1}{n+1} \right)$$

เมื่อ	UX_1	แทน	สถิติการทดสอบที่มีการแจกแจงคล้ายกับ χ^2 ที่ $df = m - 1$
	N	แทน	จำนวนนักเรียนที่ทดสอบ
	n	แทน	จำนวนค่าสังเกตหรือจำนวนข้อสอบ
	m	แทน	จำนวนค่าสัมประสิทธิ์ความเชื่อมั่น
	r_i	แทน	ค่าสัมประสิทธิ์การสรุปอ้างอิง
	S_u^2	แทน	ความแปรปรวนของคะแนน u_i
	$\bar{u}$	แทน	ค่าเฉลี่ยของคะแนน u_i
	C_u	แทน	ค่าปรับแก้ความแปรปรวนของ u_i ในกรณีที่ค่าสัมประสิทธิ์ การสรุปอ้างอิงสรุปคำนวณจากกลุ่มตัวอย่างกลุ่มเดียว
	$r(x_i, x_j)$	แทน	ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบ i และ j เมื่อใช้คะแนนรวมเป็นหน่วยการวิเคราะห์

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยดำเนินการตามลำดับดังนี้

สัญลักษณ์และอักษรย่อในการวิเคราะห์ข้อมูล

i	แทน	ผลขององค์ประกอบของข้อสอบในเอกภพของการสังเกตที่ยอมรับได้
r	แทน	ผลขององค์ประกอบของผู้ตรวจในเอกภพของการสังเกตที่ยอมรับได้
p	แทน	ผลขององค์ประกอบของผู้เข้าสอบในเอกภพของการสังเกตที่ยอมรับได้
ip	แทน	ผลร่วมระหว่างข้อสอบและผู้เข้าสอบที่กำหนด
rp	แทน	ผลร่วมระหว่างผู้ตรวจและผู้เข้าสอบที่กำหนด
ir	แทน	ผลร่วมระหว่างข้อสอบและผู้ตรวจที่กำหนด
irp	แทน	ผลร่วมระหว่างข้อสอบ ผู้ตรวจ และผู้เข้าสอบ ซึ่งปะปนอยู่กับ ความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้
$p:r$	แทน	ผลขององค์ประกอบของผู้เข้าสอบที่แฝงอยู่ในผู้ตรวจ
$ip:r$	แทน	ผลร่วมระหว่างข้อสอบกับผู้เข้าสอบที่แฝงอยู่ในผู้ตรวจ
$\bar{X}$	แทน	ค่าเฉลี่ย
$S.D.$	แทน	ความเบี่ยงเบนมาตรฐาน
$C.V.$	แทน	สัมประสิทธิ์ความผันแปร
λ_i^2	แทน	สัดส่วนระหว่างผลรวมของเมตริกในแถวที่ i กับความแปรปรวนของแบบทดสอบ
ΣS_i^2	แทน	ผลรวมของความแปรปรวนของคะแนนในแต่ละทักษะ
S_X^2	แทน	ความแปรปรวนของคะแนนของแบบทดสอบทั้งฉบับ
r_{F-R}	แทน	ความเชื่อมั่นของแบบทดสอบที่คำนวณโดยใช้สูตรของ Feldt - Raju
S_E	แทน	ความคลาดเคลื่อนมาตรฐานในการวัด
α	แทน	ผลขององค์ประกอบต่าง ๆ ที่เขียนในรูปทั่วไป แทนได้ทั้งผลหลัก (Main Effect) และผลร่วม (Interaction Effect) ในขั้นการวิเคราะห์ความแปรปรวน
df	แทน	องศาความเป็นอิสระ
SS	แทน	ผลรวมของกำลังสองของค่าเบี่ยงเบน
MS	แทน	ค่าเฉลี่ยของกำลังสองของค่าเบี่ยงเบน
σ^2	แทน	ค่าประมาณความแปรปรวน
$\sigma^2(p)$	แทน	ค่าความแปรปรวนของคะแนนเอกภพ

$\sigma^2(\delta)$ แทน ค่าความแปรปรวนของคะแนนความคลาดเคลื่อน

$E\rho^2$ แทน ค่าสัมประสิทธิ์การสุปอ้างอิง

n_r, n'_r, n'_p แทน ขนาดกลุ่มตัวอย่างขององค์ประกอบหรือระดับขององค์ประกอบข้อสอบ
ผู้ตรวจและผู้เข้าสอบ ในขั้นการวิเคราะห์ความแปรปรวน

UX_1 แทน ค่าสถิติทดสอบที่มีการแจกแจงคล้าย chi - square (χ^2)

การนำเสนอผลการวิเคราะห์ข้อมูล

การศึกษาค้นคว้าในครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับขั้นตอนดังนี้

1. ค่าสถิติพื้นฐานของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ
2. ค่าความเชื่อมั่นของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ
3. การประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสุปอ้างอิงของการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ
4. การศึกษาเพื่อตัดสินใจ (D Study) สุปอ้างอิงจำนวนผู้เข้าสอบที่ถูกสุ่มมาตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ
5. การเปรียบเทียบค่าสัมประสิทธิ์การสุปอ้างอิงของการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

ผลการวิเคราะห์ข้อมูล

ตอนที่ 1 ค่าสถิติพื้นฐานของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

การวิเคราะห์ข้อมูลในขั้นตอนนี้ ผู้วิจัยได้นำคะแนนจากแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการไปให้ผู้ตรวจ 5 คน ตรวจให้คะแนนผู้เข้าสอบทุกคนและตรวจทุกทักษะ โดยใช้กฎเกณฑ์การให้คะแนน (Scoring Rubric) ที่ผู้วิจัยสร้างขึ้นเป็นคู่มือในการตรวจ แล้วนำผลมาวิเคราะห์หาค่าสถิติพื้นฐานของแบบทดสอบ จำแนกผลตามผู้ตรวจ ปรากฏผลดังตาราง 12

ตาราง 12 ค่าสถิติพื้นฐานของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

ผู้ตรวจคนที่	จำนวนทักษะ	คะแนนเต็ม	$\bar{X}$	$S.D.$	$C.V.(%)$
1	5	20	11.120	2.309	20.77
2	5	20	10.880	2.201	20.23
3	5	20	10.940	2.113	19.32
4	5	20	10.920	2.320	21.24
5	5	20	11.360	2.238	19.70
รวม			11.044	2.202	20.25

ผลการวิเคราะห์ พบว่า สถิติพื้นฐานของคะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ มีค่าเฉลี่ยจำแนกตามผู้ตรวจคนที่ 1 – 5 เท่ากับ 11.120, 10.880, 10.940, 10.920 และ 11.360 คะแนน ตามลำดับ ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.309, 2.201, 2.113, 2.320 และ 2.238 ตามลำดับ และค่าสัมประสิทธิ์ความแปรผัน (Coefficient of Variation) เท่ากับ 20.766 %, 20.226 %, 19.316 %, 21.243 % และ 19.704 % ตามลำดับ และคะแนนทั้งฉบับของแบบทดสอบมีค่าเฉลี่ยเท่ากับ 11.044 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.202 และค่าสัมประสิทธิ์ความแปรผัน เท่ากับ 20.25 %

ตอนที่ 2 การหาค่าความเชื่อมั่นของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

บูรณาการ

ผู้วิจัยนำผลการตรวจของผู้ตรวจแต่ละคน มาคำนวณหาค่าความเชื่อมั่นของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการทั้งฉบับ ใช้สูตร r_{F-R} ของ Feldt - Raju แยกคำนวณตามผลการตรวจของผู้ตรวจแต่ละคน นำค่าความเชื่อมั่นของผู้ตรวจแต่ละคนมาหาค่าเฉลี่ย แล้วนำผลที่ได้ไปคำนวณหาค่าความเชื่อมั่นรวมของผู้ตรวจทุกคน โดยใช้สูตรปรับขยายของสเฟียร์แมน - บราวน์ (บุญเชิด ภิญโญนนตพงษ์, 2545 : 150) พร้อมกับคำนวณค่าความคลาดเคลื่อนมาตรฐานในการวัด ปรากฏผลดังตาราง 13

ตาราง 13 ค่าความเชื่อมั่นของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

ผู้ตรวจคนที่	$\Sigma \lambda_i^2$	ΣS_i^2	S_x^2	r_{F-R}	S_E
1	0.2107	1.7674	5.3322	0.8470	0.9032
2	0.2038	1.5976	4.8424	0.8416	0.8760
3	0.2054	1.5000	4.4657	0.8358	0.8562
4	0.2015	1.6759	5.3812	0.8623	0.8609
5	0.2081	1.4440	5.0106	0.8989	0.7116
รวม				0.9677	0.3957

ผลการวิเคราะห์ พบว่า ค่าความเชื่อมั่นของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ คำนวณโดยใช้สูตร r_{F-R} ของ Feldt-Raju จำแนกตามผู้ตรวจคนที่ 1-5 เท่ากับ 0.8470, 0.8416, 0.8358, 0.8623 และ 0.8989 ตามลำดับ ค่าความคลาดเคลื่อนมาตรฐานในการวัด จำแนกตามผู้ตรวจคนที่ 1-5 เท่ากับ 0.9032, 0.8760, 0.8562, 0.8609 และ 0.7116 เมื่อนำค่าความเชื่อมั่นของผู้ตรวจแต่ละคนมาหาค่าความเชื่อมั่นรวมของผู้ตรวจ โดยใช้สูตรปรับขยายของสเฟียร์แมน - บราวน์ (บุญเชิด ภิญโญนนตพงษ์, 2545 : 150) พบว่า มีค่าความเชื่อมั่น เท่ากับ 0.9677 ค่าความคลาดเคลื่อนมาตรฐานในการวัดของผู้ตรวจทุกคน เท่ากับ 0.3957

ตอนที่ 3 การประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

ผู้วิจัยได้แยกดำเนินการตามรูปแบบของการตรวจให้คะแนน ดังนี้

3.1 การประมาณค่าความแปรปรวนเมื่อความสัมพันธ์ขององค์ประกอบเป็น $i \times r \times p$ ซึ่งจัดเป็นรูปแบบการตรวจให้คะแนนที่ผู้ตรวจทุกคน ตรวจให้คะแนนแบบทดสอบทุกทักษะของผู้เข้าสอบทุกคน การวิเคราะห์ข้อมูลในขั้นตอนนี้ ผู้วิจัยมีจุดมุ่งหมายเพื่อคำนวณหาขนาดความแปรปรวนจากแหล่งต่าง ๆ ในเอกภพค่าสังเกตที่เป็นไปได้ทั้งหมด ที่แยกเป็นอิสระจากกันของ 3 องค์ประกอบ ได้แก่ ข้อสอบ (i) ผู้ตรวจ (r) และ ผู้เข้าสอบ (p) รวมทั้งผลร่วมระหว่าง ข้อสอบกับผู้ตรวจ (ir) ข้อสอบกับผู้เข้าสอบ (ip) ผู้ตรวจกับผู้เข้าสอบ (rp) และผลร่วมระหว่างข้อสอบ ผู้ตรวจ และผู้เข้าสอบ (irp) ซึ่งปะปนอยู่กับความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ ผลการตรวจให้คะแนนและวิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป GENOVA ของ Crick และ Brennan ปรากฏผลดังตาราง 14

ตาราง 14 ค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวัด

ทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ เมื่อความสัมพันธ์ขององค์ประกอบเป็น $i \times r \times p$
($n_i = 5$, $n'_r = 5$, $n'_p = 50$)

แหล่งความแปรปรวน	$df(\alpha)$	$SS(\alpha)$	$MS(\alpha)$	$\sigma^2(\alpha)$	ร้อยละ
i	4	61.7680	15.4420	0.0436	11.46
r	4	8.9200	2.2300	0.0000	0.00
p	49	38.5600	0.7869	0.0265	6.97
ir	16	72.8320	4.5520	0.0866	22.77
ip	196	39.9920	0.2040	0.0000	0.000
rp	196	28.4400	0.1451	0.0000	0.000
irp	784	175.4080	0.2237	0.2237	58.81
รวม	1249	425.9200		0.3804	100.00

ผลการวิเคราะห์ พบว่า ผลการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ตามแบบแผน $i \times r \times p$ ค่าความแปรปรวนของแหล่งต่าง ๆ เมื่อเทียบกับค่าความแปรปรวนรวม มีค่าเรียงลำดับจากน้อยไปหามาก ดังนี้ ความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจน $\sigma^2(irp)$ คิดเป็นร้อยละ 58.807 รองลงมา คือ ค่าความแปรปรวนร่วมระหว่าง ผู้ตรวจและข้อสอบ $\sigma^2(ir)$ คิดเป็นร้อยละ 22.766 ค่าความแปรปรวนของข้อสอบ $\sigma^2(i)$ คิดเป็นร้อยละ 11.467 ค่าความแปรปรวนของผู้เข้าสอบ $\sigma^2(p)$ คิดเป็นร้อยละ 6.966 ส่วนค่าความแปรปรวนของผู้ตรวจ $\sigma^2(r)$

ค่าความแปรปรวนร่วมระหว่างข้อสอบและผู้เข้าสอบ $\sigma^2(ip)$ และค่าความแปรปรวนร่วมระหว่าง ผู้ตรวจ และผู้เข้าสอบ $\sigma^2(rp)$ มีค่าโดยประมาณไม่แตกต่างกัน

3.2 การประมาณค่าความแปรปรวนเมื่อความสัมพันธ์ขององค์ประกอบเป็น $i \times (p:r)$ ซึ่งจัดเป็นรูปแบบการตรวจให้คะแนนที่ผู้ตรวจทุกคน ตรวจให้คะแนนแบบทดสอบทุกทักษะ ของผู้เข้าสอบบางคนโดยที่ไม่ซ้ำกัน การวิเคราะห์ข้อมูลในขั้นตอนนี้ ผู้วิจัยมีจุดมุ่งหมายเพื่อคำนวณหาขนาดความแปรปรวนจากแหล่งต่าง ๆ ในเอกภพค่าสังเกตที่เป็นไปได้ทั้งหมด ที่แยกเป็นอิสระจากกันของ 3 องค์ประกอบ ได้แก่ ข้อสอบ (i) ผู้ตรวจ (r) และผู้เข้าสอบ (p) รวมทั้งผลขององค์ประกอบของผู้เข้าสอบที่แฝงอยู่ในผู้ตรวจ ($p:r$) ผลร่วมระหว่างข้อสอบกับผู้ตรวจ (ir) และผลร่วมระหว่างข้อสอบกับผู้เข้าสอบที่แฝงอยู่ในผู้ตรวจ ($ip:r$) ผลการตรวจให้คะแนนและวิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป GENOVA ของ Crick และ Brennan ปรากฏผลดัง ตาราง 15

ตาราง 15 ค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวัด

ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ เมื่อความสัมพันธ์ขององค์ประกอบเป็น $i \times (p:r)$
($n_i = 5$, $n'_r = 5$, $n'_p = 50$)

แหล่งความแปรปรวน	$df(\alpha)$	$SS(\alpha)$	$MS(\alpha)$	$\sigma^2(\alpha)$	ร้อยละ
i	4	4.1200	1.0300	0.0133	5.08
r	4	3.0000	0.7500	0.0049	1.87
$p:r$	45	15.1000	0.3356	0.0271	10.34
ir	16	5.8800	0.3675	0.0168	6.41
$ip:r$	180	36.0000	0.2000	0.2000	76.30
รวม	249	61.1000		0.2621	100.00

ผลการวิเคราะห์ พบว่า ผลการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ โดยให้ผู้ตรวจแต่ละคนสุ่มผู้เข้าสอบจำนวน 10 คน มาตรวจให้คะแนน ตามแบบแผน $i \times (p:r)$ ค่าความแปรปรวนของแหล่งต่าง ๆ เมื่อเปรียบเทียบกับค่าความแปรปรวนรวม มีค่าเรียงลำดับจากน้อยไปหามาก ดังนี้ ความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจน $\sigma^2(ip:r)$ คิดเป็นร้อยละ 76.30 รองลงมา คือ ความแปรปรวนร่วมของผู้เข้าสอบที่แฝงอยู่ในผู้ตรวจ $\sigma^2(p:r)$ คิดเป็นร้อยละ 10.34 ความแปรปรวนร่วมระหว่างข้อสอบและผู้ตรวจ $\sigma^2(ir)$ คิดเป็นร้อยละ 6.41 ความแปรปรวนของข้อสอบ $\sigma^2(i)$ คิดเป็นร้อยละ 5.08 ส่วนความแปรปรวนของผู้ตรวจ $\sigma^2(r)$ คิดเป็นร้อยละ 1.87 ซึ่งถือว่ามีค่าน้อยที่สุด

ตอนที่ 4 การศึกษาเพื่อตัดสินใจ (D Study) สรุปอ้างอิงจำนวนผู้เข้าสอบที่ถูกสุ่มมาตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่มีผู้ตรวจจำนวน 5 คน

การวิเคราะห์ข้อมูลในขั้นตอนนี้ เป็นการศึกษาเพิ่มเติมโดยใช้ข้อมูลจากขั้น G Study มาใช้สรุปอ้างอิงค่าความเชื่อมั่นของรูปแบบการตรวจให้คะแนนทั้งสองรูปแบบ และตัดสินใจเลือกจำนวนผู้เข้าสอบที่ถูกสุ่มมาตรวจคะแนนในจำนวนที่เหมาะสม เพื่อลดความคลาดเคลื่อนและเพิ่มความแม่นยำในการสรุปอ้างอิง ปรากฏผลดังตาราง 16 – 17

ตาราง 16 ค่าความแปรปรวนของคะแนนเอกภพ คะแนนความคลาดเคลื่อน และสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ตรวจให้คะแนนโดยใช้แบบแผน $i \times r \times p$ ($n_i = 5$, $n_r = 5$, $n_p = 50$)

จำนวนผู้เข้าสอบ	$\sigma^2(\rho)$	$\sigma^2(\delta)$	$E\rho^2$
10	.04356	.02179	.6666
20	.04356	.01955	.6902
30	.04356	.01880	.6985
40	.04356	.01843	.7027
50	.04356	.01821	.7052

ผลการวิเคราะห์พบว่า การตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ที่มี 5 ทักษะ ผู้ตรวจ 5 คน และผู้เข้าสอบจำนวน 10, 20, 30, 40 และ 50 คน ค่าสัมประสิทธิ์การสรุปอ้างอิงเท่ากับ .6666, .6902, .6984, .7027 และ .7052 ตามลำดับ

ตาราง 17 ค่าความแปรปรวนของคะแนนเอกภพ คะแนนความคลาดเคลื่อน และสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ตรวจให้คะแนนโดยใช้แบบแผน $i \times (p:r)$ ($n_i = 5$, $n_r = 5$, $n_p = 50$)

จำนวนผู้เข้าสอบ	$\sigma^2(\rho)$	$\sigma^2(\delta)$	$E\rho^2$
10	.01325	.00735	.6432
20	.01325	.00535	.7124
30	.01325	.00468	.7389
40	.01325	.00435	.7528
50	.01325	.00415	.7615

ผลการวิเคราะห์พบว่า การตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ที่มี 5 ทักษะ ผู้ตรวจ 5 คน และผู้เข้าสอบจำนวน 50 คน แต่ถูกสุ่มมาตรวจให้คะแนนที่จำนวน 10, 20, 30, 40 และ 50 คน ค่าสัมประสิทธิ์การสุ่มอ้างอิง เท่ากับ .6432, .7124, .7389, .7528 และ .7615 ตามลำดับ

ตอนที่ 5 การเปรียบเทียบค่าสัมประสิทธิ์การสุ่มอ้างอิงของการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

การวิเคราะห์ตามขั้นตอนนี้ มีจุดมุ่งหมายสำคัญ เพื่อเปรียบเทียบค่าสัมประสิทธิ์การสุ่มอ้างอิงของวิธีการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ทั้งสองรูปแบบ ได้แก่ รูปแบบการตรวจให้คะแนนที่กำหนดให้ผู้ตรวจทุกคน ($n_r = 5$) ตรวจให้คะแนนผู้เข้าสอบทุกคน ($n_p = 50$) และตรวจทุกทักษะ ($n_i = 5$) ประมาณค่าสัมประสิทธิ์การสุ่มอ้างอิงได้เท่ากับ .7052 และรูปแบบการตรวจให้คะแนนที่กำหนดให้ผู้ตรวจทุกคน ($n_r = 5$) ตรวจให้คะแนนผู้เข้าสอบที่จัดเป็นกลุ่ม ๆ ละ 10 คน โดยการสุ่มอย่างง่าย (Simple Random Sampling) ผู้ตรวจ 5 คน แยกกันตรวจคนละ 1 กลุ่มและตรวจทุกทักษะ ($n_i = 5$) ประมาณค่าสัมประสิทธิ์การสุ่มอ้างอิงได้เท่ากับ .6432 เปรียบเทียบค่าสัมประสิทธิ์การสุ่มอ้างอิงของรูปแบบการตรวจให้คะแนนทั้งสองรูปแบบ ที่คำนวณด้วยสูตร UX_1 ของ Woodruff and Feldt. ปรากฏผลดังตาราง 18

ตาราง 18 การเปรียบเทียบค่าสัมประสิทธิ์การสุ่มอ้างอิงของการตรวจให้คะแนนแบบทดสอบทักษะ

กระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ตรวจให้คะแนนโดยใช้แบบแผน $i \times r \times p$ ($n_i = 5$, $n_r = 5$, $n_p = 50$) และแบบแผน $i \times (p : r)$ ($n_i = 5$, $n_r = 5$, $n_p = 50$)

เอกภพของการสุ่มอ้างอิง	$\sigma^2(\rho)$	$\sigma^2(\delta)$	$E\rho^2$	UX_1	p
$i \times r \times p$	0.0436	0.0182	0.7052	25.500	.00002
$i \times (p : r)$	0.0133	0.0074	0.6432		

ผลการวิเคราะห์ พบว่า ค่าสัมประสิทธิ์การสุ่มอ้างอิงของการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ตรวจให้คะแนนตามแบบแผน $i \times r \times p$ มีค่าเท่ากับ 0.7052 และค่าสัมประสิทธิ์การสุ่มอ้างอิงของการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ที่ตรวจให้คะแนนตามแบบแผน $i \times (p : r)$ มีค่าเท่ากับ 0.6432 จากการทดสอบความแตกต่างด้วย UX_1 ได้ค่าเท่ากับ 25.50 ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ .01 ผลการวิเคราะห์ในขั้นตอนนี้เป็นเป้าหมายหลักในการวิจัย ผู้วิจัยจึงได้นำเอาผลการวิเคราะห์ข้อมูลในขั้นตอนนี้ไปใช้สรุปความมุ่งหมายของการวิจัยในข้อที่ 2 และทดสอบสมมุติฐานของการวิจัย

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

ความมุ่งหมายของการวิจัย

1. เพื่อพัฒนาเกณฑ์การให้คะแนน (Scoring Rubric) ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ
2. เพื่อศึกษารูปแบบของการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการสองรูปแบบที่คำนวณโดยใช้ทฤษฎีการสรุปอ้างอิง

สมมุติฐานของการวิจัย

ความเชื่อมั่นของการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่มีรูปแบบการตรวจต่างกัน ให้ค่าแตกต่างกัน

ประชากรที่ใช้ในการวิจัย

ประชากรที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนมัธยมศึกษาสังกัดกรมสามัญศึกษาจังหวัดนครราชสีมา ปีการศึกษา 2544

กลุ่มตัวอย่างที่ใช้ในการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัย เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 4 โรงเรียนบุญวัฒนา 2 สังกัดกรมสามัญศึกษาจังหวัดนครราชสีมา ปีการศึกษา 2544 จำนวน 50 คน ซึ่งเลือกมาโดยการสุ่มแบบหลายขั้นตอน (Multi - Stage Random Sampling)

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการศึกษาค้นคว้าในครั้งนี้ เป็นแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ มีลักษณะเป็นคำถามปลายเปิดและกำหนดสถานการณ์ที่เกี่ยวกับวิทยาศาสตร์ให้ตอบประกอบด้วย 5 ทักษะ คือ

- | | |
|---|-------------|
| 1. ทักษะการตั้งสมมุติฐาน | จำนวน 3 ข้อ |
| 2. ทักษะการกำหนดนิยามเชิงปฏิบัติการ | จำนวน 3 ข้อ |
| 3. ทักษะการกำหนดและควบคุมตัวแปร | จำนวน 4 ข้อ |
| 4. ทักษะการทดลอง | จำนวน 1 ข้อ |
| 5. ทักษะการตีความหมายข้อมูลและลงข้อสรุป | จำนวน 4 ข้อ |

รวมเป็นแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ 1 ฉบับ จำนวน 5 ทักษะ จำนวนข้อสอบรวม 15 ข้อ ใช้เวลา 120 นาที

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลในการวิจัย ดังนี้

1. นำหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อขอความอนุเคราะห์กับโรงเรียนที่เป็นกลุ่มตัวอย่าง
2. ติดต่อประสานงาน ขออนุญาตผู้บริหารและครูผู้สอนของโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลา สถานที่ วิธีดำเนินการสอบ
3. นำแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผ่านการวิเคราะห์คุณภาพรายข้อแล้ว ไปทดสอบกับกลุ่มตัวอย่าง จำนวน 50 คน ผู้ดำเนินการสอบ 2 คน ซึ่งผู้ดำเนินการสอบชี้แจงการปฏิบัติในการดำเนินการสอบ ดังนี้
 - 3.1 อธิบายให้นักเรียนกลุ่มตัวอย่างเข้าใจวัตถุประสงค์ และผลประโยชน์ที่จะได้รับจากการทำแบบทดสอบ
 - 3.2 อธิบายให้ผู้เข้าสอบทุกคนเข้าใจถึงวิธีดำเนินการสอบ วิธีการปฏิบัติ และวิธีตอบก่อนที่จะให้ผู้เข้าสอบลงมือทำ
4. นำกระดาษคำตอบของนักเรียนทั้งหมดมาถ่ายเอกสารจำนวน 6 ชุด จัดให้ผู้ตรวจจำนวน 5 คน ได้ตรวจให้คะแนนตามแนวทางในการตอบและกฎเกณฑ์การให้คะแนนที่ผู้วิจัยได้สร้างขึ้นเป็นคู่มือการตรวจให้คะแนน โดยแบ่งการดำเนินการเป็น 2 วิธีดังนี้
 - 4.1 นำกระดาษคำตอบชุดที่ 1 มาจัดแบ่งออกเป็น 5 กลุ่ม ๆ ละ 10 คน นำไปให้ผู้ตรวจให้คะแนนแยกกันตรวจ 1 คน ต่อ 1 กลุ่ม (ตามรูปแบบผู้ตรวจทุกคนตรวจให้คะแนนทุกทักษะของผู้เข้าสอบบางคน) แล้วบันทึกผลการตรวจให้คะแนน
 - 4.2 หลังจากนั้น 1 สัปดาห์ นำกระดาษคำตอบจำนวน 5 ชุด ที่เหลือไปให้ผู้ตรวจทุกคนได้ตรวจให้คะแนนคนละ 1 ชุด (ตามรูปแบบผู้ตรวจทุกคนตรวจให้คะแนนทุกทักษะของผู้เข้าสอบทุกคน) แล้วบันทึกผลการตรวจให้คะแนน
5. นำผลที่ได้จากการตรวจให้คะแนนมาวิเคราะห์ข้อมูล ดังนี้
 - 5.1 คำนวณค่าสถิติพื้นฐานของแบบทดสอบจำแนกตามผู้ตรวจให้คะแนน ตามข้อ 4.2
 - 5.2 คำนวณค่าความเชื่อมั่นของแบบทดสอบ โดยใช้สูตร r_{F-R} ของ Feldt-Raju (บุญเชิด ภิญโญนนตพงษ์, 2545 : 139)
 - 5.3 คำนวณค่าสัมประสิทธิ์การสรุปอ้างอิงของรูปแบบการตรวจให้คะแนน จากสูตรของ เบรนนอน วิเคราะห์โดยใช้โปรแกรมสำเร็จรูป GENOVA โดยมีรูปแบบในการคำนวณ ดังนี้
 - 5.3.1 นำผลการตรวจให้คะแนนในข้อ 4.2 มาคำนวณสัมประสิทธิ์การสรุปอ้างอิงตามแบบแผน $i \times r \times p$ ซึ่งจัดเป็นรูปแบบ Two – Facet Crossed Design
 - 5.3.2 นำผลการตรวจให้คะแนนในข้อ 4.1 มาคำนวณสัมประสิทธิ์การสรุปอ้างอิงตามแบบแผน $i \times (p : r)$ ซึ่งจัดเป็นรูปแบบ Two – Facet Nested Design

5.4. ทดสอบความแตกต่างของผลการตรวจให้คะแนนทั้งสองรูปแบบ (ข้อ 5.3.1 และ 5.3.2) ตามสมมติฐานของการวิจัย โดยใช้สูตร UX_1 ของวูดรUFF และเฟลด์ (Woodruff and Feldt. 1986 : 393-413)

สรุปผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานและค่าความเชื่อมั่นของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

ค่าสถิติพื้นฐานของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ เมื่อพิจารณาผลการตรวจให้คะแนนจากผู้ตรวจทุกคน มีค่าเฉลี่ยจำแนกตามผู้ตรวจคนที่ 1 – 5 เท่ากับ 11.120, 10.880, 10.940, 10.920 และ 11.360 คะแนน ตามลำดับ มีค่าเฉลี่ยรวม เท่ากับ 11.044 คะแนน เมื่อพิจารณาค่าเฉลี่ยจำแนกตามผู้ตรวจและค่าเฉลี่ยรวม พบว่า มีผลการตรวจที่สอดคล้องกันทั้ง 5 คน และมีค่าใกล้เคียงกับค่าเฉลี่ยรวม ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.202 และค่าสัมประสิทธิ์ความแปรผัน เท่ากับ 20.25 % พบว่า มีค่าที่ต่ำมาก แสดงว่า การตรวจให้คะแนนมีการเบี่ยงเบนบ้างเล็กน้อย คะแนนที่ได้จากการตรวจเกาะกลุ่มไม่กระจกระบาย

ค่าความเชื่อมั่นที่คำนวณโดยใช้สูตร r_{F-R} ของ Feldt – Raju จำแนกตามผู้ตรวจคนที่ 1 – 5 มีค่าเท่ากับ .8470, .8416, .8358, .8623 และ .8989 ตามลำดับ เมื่อนำความเชื่อมั่นของผู้ตรวจแต่ละคนมาหาค่าเฉลี่ยรวมของความเชื่อมั่นเท่ากับ .8571 นำค่าเฉลี่ยรวมไปปรับขยายโดยใช้สูตรของสเปียร์แมน – บราวน์ ได้ค่าความเชื่อมั่นของผู้ตรวจทุกคน เท่ากับ .9677 แสดงถึงคุณภาพและประสิทธิภาพของแบบทดสอบและกฎเกณฑ์การให้คะแนนที่ผู้วิจัยสร้างขึ้น ที่เป็นองค์ประกอบอยู่ภายในแบบทดสอบ และหลักฐานที่แสดงว่ามีความน่าเชื่อถือมาก ก็คือ ความคลาดเคลื่อนมาตรฐานในการวัด ซึ่งมีค่าเท่ากับ .3957 ถือว่าน้อยมาก

2. จากการศึกษาแบบการตรวจให้คะแนน 2 รูปแบบ โดยอาศัยการประมาณค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง คำนวณโดยใช้ทฤษฎีการสรุปอ้างอิง และทดสอบความแตกต่างของทั้งสองรูปแบบด้วยสูตร UX_1 ของ Woodruff and Feldt. ได้ผลดังนี้

การตรวจให้คะแนนตามแบบแผน $i \times r \times p$ พบว่า ความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจน $\sigma^2(irp)$ มีค่ามากที่สุด ค่าความแปรปรวนของผู้ตรวจ $\sigma^2(r)$ ค่าความแปรปรวนร่วมระหว่างข้อสอบและผู้เข้าสอบ $\sigma^2(ip)$ และค่าความแปรปรวนร่วมระหว่างผู้ตรวจและผู้เข้าสอบ $\sigma^2(rp)$ มีค่าน้อยที่สุด ส่วนการตรวจให้คะแนนตามแบบแผน $i \times (p : r)$ พบว่า ความแปรปรวนของความคลาดเคลื่อนเชิงสุ่มที่ไม่สามารถระบุแหล่งได้ชัดเจน $\sigma^2(ip : r)$ มีค่ามากที่สุด ความแปรปรวนของผู้ตรวจ $\sigma^2(r)$ มีค่าน้อยที่สุด

ค่าสัมประสิทธิ์การสรุปอ้างอิง ผลการตรวจแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ ที่คำนวณโดยใช้แบบแผน $i \times r \times p$ เท่ากับ .7052 และค่าสัมประสิทธิ์การสรุปอ้างอิงผลการตรวจแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ ที่คำนวณโดยใช้แบบแผน $i \times (p : r)$ เท่ากับ .6432 และเมื่อทำการทดสอบนัยสำคัญของความแตกต่างของค่าความเชื่อมั่นโดยใช้สูตร UX_1 ของ Woodruff and Feldt. พบว่า มีค่าแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ .01

อภิปรายผล

1. กฎเกณฑ์การให้คะแนน (Scoring Rubric) ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผู้วิจัยสร้างขึ้น สามารถพัฒนาให้มีคุณภาพและประสิทธิภาพ พิจารณาผลจากค่าสถิติพื้นฐาน ความเชื่อมั่นของแบบทดสอบและการวิเคราะห์ค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงดังต่อไปนี้

ค่าเฉลี่ยจากผลการตรวจของผู้ตรวจทั้ง 5 คน และค่าเฉลี่ยรวมทั้งฉบับมีค่าใกล้เคียงกัน แสดงว่าข้อวินิจฉัยในการให้คะแนนผู้เข้าสอบของผู้ตรวจแต่ละคนมีความสอดคล้องกัน ค่าส่วนเบี่ยงเบนมาตรฐานรวมและค่าสัมประสิทธิ์ความแปรผันของแบบทดสอบทั้งฉบับอยู่ในระดับต่ำ แสดงว่า คะแนนของผู้ตรวจแต่ละคนเกาะกลุ่มกันไม่กระจัดกระจายและมีการเบี่ยงเบนของคะแนนน้อยมาก เมื่อพิจารณาที่ค่าความแปรปรวนขององค์ประกอบที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง พบว่า ค่าความแปรปรวนของผู้ตรวจทั้งสองรูปแบบมีค่าน้อยมาก เนื่องจาก ผู้ตรวจทุกคนได้ใช้กฎเกณฑ์การให้คะแนนที่ผู้วิจัยสร้างขึ้นเป็นคู่มือในการตรวจ จึงทำให้อคติของผู้ตรวจจะถูกจัดออกไปอย่างสิ้นเชิง แสดงถึงความเป็นปรนัยของกฎเกณฑ์การให้คะแนนที่ผู้วิจัยสร้างขึ้น ช่วยให้ผู้ตรวจมีความเห็นที่สอดคล้องกัน จึงไม่พบว่ามีค่าความแปรปรวนระหว่างผู้ตรวจเกิดขึ้น การวิจัยครั้งนี้มีผลการศึกษาที่แตกต่างจาก อิบราฮิม (Ibrahim, 1984 : 499-A) ที่พบว่า ในการประเมินโดยใช้ผู้เชี่ยวชาญนั้น ผู้เชี่ยวชาญเป็นแหล่งความแปรปรวนที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิงมากที่สุด และแตกต่างจาก ไพร์ธน์ วงษ์นาม (2533 : 112) ที่พบว่า ผู้ตรวจเป็นแหล่งความแปรปรวนที่มีผลต่อการสรุปอ้างอิงของแบบทดสอบความเรียงมากที่สุด ความแปรปรวนของผู้เข้าสอบมีผลรองลงมา แต่สอดคล้องกับ ชูชาน เลน และคณะ (Lane, S. and others, 1969 : 71-79) ที่พบว่า ความคลาดเคลื่อนเนื่องมาจากผู้ตรวจจะมีค่าน้อยมาก เมื่อมีการตรวจให้คะแนนแบบประเมินรวมและมีเกณฑ์การให้คะแนน ดังนั้น เพื่อให้ผลการตรวจให้คะแนนเกิดความสอดคล้องภายในผู้ตรวจสูงขึ้น ควรมีการออกแบบกฎเกณฑ์การให้คะแนนและการฝึกหรืออบรมให้กับผู้ตรวจ แต่งานวิจัยนี้ผู้วิจัยเลือกวิธีการพัฒนากฎเกณฑ์การให้คะแนนให้มีคุณภาพ และจากผลการวิจัยจึงสรุปได้ว่า กฎเกณฑ์การให้คะแนน (Scoring Rubric) ทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่ผู้วิจัยสร้างขึ้น สามารถพัฒนาให้มีคุณภาพและประสิทธิภาพ เป็นไปตามความมุ่งหมายของการวิจัยทุกประการ

2. การศึกษารูปแบบการตรวจให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการที่คำนวณโดยใช้ทฤษฎีการสรุปอ้างอิง

ค่าสัมประสิทธิ์การสรุปอ้างอิงของการตรวจให้คะแนนตามแบบแผน $i \times r \times p$ ที่กำหนดให้ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบทุกคน มีค่าเท่ากับ .7052 และค่าสัมประสิทธิ์การสรุปอ้างอิงของการตรวจให้คะแนนตามแบบแผน $i \times (p : r)$ ที่กำหนดให้ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบเป็นรายกลุ่ม ๆ ละ 10 คน โดยไม่ซ้ำกัน มีค่าเท่ากับ .6432 เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นของการตรวจให้คะแนนทั้งสองรูปแบบโดยใช้สูตร UX_1 พบว่า แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ .01 สาเหตุที่การตรวจให้คะแนนตามแบบแผน $i \times (p : r)$ ค่าสัมประสิทธิ์การสรุปอ้างอิงจะมีค่าต่ำกว่าการตรวจให้คะแนนตามแบบแผน $i \times r \times p$ พบว่า เกิดจากการกระจายของคะแนน ถ้าแบบแผนใดมีการกระจายของคะแนนของผู้เข้าสอบมาก ค่าสัมประสิทธิ์การสรุปอ้างอิงจะมีค่าสูงขึ้น ดังที่ กรอนลันด์ (Gronlund, 1976 : 117) กล่าวว่า ถ้าความสามารถของผู้เข้าสอบในกลุ่มแตกต่างกันมาก การกระจายของคะแนนจะมากทำให้

ความเชื่อมั่นสูง ถ้าความสามารถของกลุ่มผู้เข้าสอบใกล้เคียงกันมาก การกระจายของคะแนนมีค่าน้อย ทำให้ความเชื่อมั่นต่ำ สอดคล้องกับ คันทิงแฮม (Cunningham. 1986 : 112-118) ที่พบว่า ความสามารถผู้เข้าสอบที่หลากหลายจะส่งผลให้ค่าความเชื่อมั่นสูงขึ้นเมื่อความแปรปรวนของคะแนนเพิ่มขึ้น ในแบบแผน $i \times (p : r)$ การสุ่มผู้เข้าสอบเพื่อจัดให้ผู้ตรวจได้ตรวจโดยที่ไม่ซ้ำกันนั้น ใช้วิธีการสุ่มอย่างง่ายโดยการจับฉลาก จึงอาจทำให้ผู้ตรวจบางคนได้ผลคะแนนของผู้เข้าสอบไม่กระจาย อาจสุ่มได้ผู้เข้าสอบที่มีความสามารถใกล้เคียงกัน จึงทำให้ค่าความเชื่อมั่นต่ำลง สรุปได้ว่า รูปแบบการตรวจให้คะแนนตามแบบแผน $i \times (p : r)$ ที่กำหนดให้ผู้ตรวจทุกคนตรวจแบบทดสอบทุกทักษะของผู้เข้าสอบเป็นรายกลุ่ม ๆ ละ 10 คน โดยไม่ซ้ำกันนั้น จะไม่สามารถทดแทนวิธีการที่ผู้ตรวจได้ตรวจให้คะแนนผู้เข้าสอบทุกคน ถึงแม้ว่า รูปแบบการตรวจให้คะแนนทั้งสอง จะให้ค่าความเชื่อมั่นใกล้เคียงกัน ผลการวิจัยนี้จึงเป็นไปตามสมมุติฐานที่ตั้งไว้ทุกประการ

ข้อค้นพบบางประการคือ งานวิจัยนี้มีการศึกษาผลของการตรวจให้คะแนน 2 รูปแบบ คือ รูปแบบที่ผู้ตรวจทุกคนตรวจให้คะแนนทุกทักษะของผู้เข้าสอบทุกคน และรูปแบบที่ผู้ตรวจทุกคนตรวจให้คะแนนทุกทักษะของผู้เข้าสอบบางคน โดยไม่ซ้ำกัน ในรูปแบบหลังนี้จึงได้แบ่งผู้เข้าสอบ 10 คน ต่อผู้ตรวจ 1 คน ภายใต้เงื่อนไขที่ต้องไม่ให้มีการตรวจผู้เข้าสอบซ้ำกัน ได้ผลดังที่อภิปรายผลไปแล้ว แต่จากการศึกษาเพิ่มเติมในชั้น D Study เพื่อตัดสินใจหารูปแบบที่เหมาะสม พบว่า ในรูปแบบหลังเมื่อสุ่มผู้เข้าสอบมากขึ้น คือ 20, 30, 40 และ 50 คน ค่าสัมประสิทธิ์การสุ่อ้างอิงจะมีค่าสูงขึ้นและสูงกว่าในรูปแบบที่ 1 แสดงให้เห็นว่า ถ้าสุ่มผู้เข้าสอบมาตรวจให้คะแนนที่มีจำนวนมากพอ คือ ตั้งแต่ 20 คนขึ้นไป ก็จะสามารถเพิ่มค่าสัมประสิทธิ์การสุ่อ้างอิงให้สูงขึ้นมากกว่ารูปแบบที่ 1 ได้ สรุปได้ว่า ถ้าศึกษาเพิ่มเติมการตรวจให้คะแนนในรูปแบบที่ 2 คือ รูปแบบที่ผู้ตรวจทุกคนตรวจให้คะแนนทุกทักษะของผู้เข้าสอบบางคนจะช่วยลดภาระในการตรวจให้คะแนนของผู้ตรวจได้

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1 คุณสมบัติของผู้ตรวจให้คะแนนควรกำหนดให้ชัดเจน ต้องเป็นผู้สอนในรายวิชานั้น ๆ ควรมีความเชี่ยวชาญและมีประสบการณ์ในการสอนมากพอสมควร เพราะทักษะการทดลองจำเป็นต้องใช้ประสบการณ์ตรงนี้ร่วมด้วย

1.2 เมื่อครูผู้สอนได้ดำเนินการวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการของผู้เรียนและทราบผลของคำตอบเรียบร้อยแล้ว ควรนำผลไปดำเนินการปรับปรุง / แก้ไข และพัฒนาให้เกิดพฤติกรรมด้านทักษะกระบวนการทางวิทยาศาสตร์ในตัวผู้เรียนต่อไป

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

ควรมีการศึกษาในลักษณะเดียวกันนี้ โดยการออกแบบเครื่องมือที่เน้นภาคปฏิบัติในสถานการณ์จริง หรือจำลองสถานการณ์ในการปฏิบัติ เพื่อให้เกิดความสมบูรณ์ในการประเมินพฤติกรรมที่แสดงถึงความสามารถด้านทักษะกระบวนการทางวิทยาศาสตร์อย่างแท้จริง

บรรณานุกรม

บรรณานุกรม

- กิ่งฟ้า สันฐวณิช และคณะ. (2525). ผลการวิจัยเรื่องทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนระดับมัธยมศึกษา. ขอนแก่น : มหาวิทยาลัยขอนแก่น.
- กนกวรรณ เอี่ยมชัย. (2539). การศึกษาค่าความเชื่อมั่นของแบบทดสอบ เอ็ม อี คิว วิชาแนวคิดพื้นฐานและหลักการพยาบาลที่ตรวจให้คะแนนต่างกัน. ปรินุณานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- กรมวิชาการ. (2533). หลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง พ.ศ. 2533). กรุงเทพฯ : กรมฯ.
- _____. (2539). การประเมินผลจากสภาพจริง. กรุงเทพฯ : กรมฯ.
- โกวิท ประวาลพฤกษ์ และ สมศักดิ์ สันธะเวชญ์. (2525). การประเมินในชั้นเรียน. กรุงเทพฯ : วัฒนาพานิช.
- กันยารัตน์ ฤทธิ์บำรุง. (2531). การศึกษาผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ทักษะกระบวนการทางวิทยาศาสตร์และความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่เรียนโดยการฝึกอภิปรายแก้ปัญหาพร้อมกันทั้งห้องเป็นกลุ่มย่อยและแก้ปัญหาเป็นรายบุคคล. ปรินุณานิพนธ์ กศ.ม. (วิทยาศาสตร์ศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- กาญจนา ลากรวย. (2532). การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ผลสัมฤทธิ์ด้านทักษะกระบวนการทางวิทยาศาสตร์และความสามารถในการแก้ปัญหาเชิงวิทยาศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 5 ที่มีระดับความสามารถทางการเรียนวิชาวิทยาศาสตร์ต่างกันโดยการสอนสาธิตแบบไม่ชี้แนวทางและการสอนสาธิตแบบชี้แนวทาง. ปรินุณานิพนธ์ กศ.ม. (วิทยาศาสตร์ศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ก้องศักดิ์ ศรีน้อย. (2527). การศึกษาการใช้คำถามเน้นทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการในการสอนแบบสืบเสาะหาความรู้ที่มีผลต่อความคิดสร้างสรรค์ทางวิทยาศาสตร์และผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3. ปรินุณานิพนธ์ กศ.ม. (วิทยาศาสตร์ศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- โครงการพัฒนาทรัพยากรมนุษย์. (ม.ป.ป.). การประเมินผลโดยใช้แฟ้มสะสมงานดีเด่น. กระบี่ : โครงการพัฒนาทรัพยากรมนุษย์ กระทรวงศึกษาธิการ.
- จิตติพร คิริตานนท์. (2542). การศึกษาผลการประเมินด้วยพอดโฟรีโอที่มีต่อความรับผิดชอบเจตคติต่อวิชาวิทยาศาสตร์และทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 2. ปรินุณานิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- จิราพร อัครสมพงษ์. (2536). การเปรียบเทียบความสอดคล้องของคะแนนแบบทดสอบความเรียงจากผู้ตรวจและวิธีการตรวจแตกต่างกัน. วิทยานิพนธ์ ค.ม. (วิจัยการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.

ชัยฤทธิ์ คีลาเดช. (2540). *การพัฒนาแฟ้มสะสมงานในการประเมินผลการเรียนวิชาภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่ 3*. ปรินญาณินพนธ์ กศ.ด. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

ชุติญา สุริยมณฑล. (2535). *การสำรวจประเภททักษะกระบวนการทางวิทยาศาสตร์ในหนังสือเรียนวิทยาศาสตร์ระดับมัธยมศึกษาตอนต้น*. วิทยานิพนธ์ ค.ม. (การมัธยมศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.

✓ ชาญ วชิรเดช. (2538). *สัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบการแก้ปัญหาที่สอบด้วยไมโครคอมพิวเตอร์*. ปรินญาณินพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

จำเริญ วงศ์สว่าง. (2528). *ผลการทดสอบย่อยโดยใช้แบบทดสอบเลือกตอบ เดิมคำ และเลือกตอบที่มีต่อความสามารถในการเขียนภาษาไทย ของนักเรียนชั้นมัธยมศึกษาปีที่ 3*. ปรินญาณินพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

✓ แดง กลางท่าไผ่. (2533). *การประยุกต์ทฤษฎีการสรุปอ้างอิงในการหาความเชื่อมั่นของการประเมินความตรงเชิงเนื้อหา*. วิทยานิพนธ์ ศศ.ม. (การวัดและประเมินผลการศึกษา). ขอนแก่น : บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น. ถ่ายเอกสาร.

ญานี ทองทับ. (2536). *การศึกษาทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 6 ในโรงเรียนสังกัดกรุงเทพมหานคร*. ปรินญาณินพนธ์ กศ.ม. (วิทยาศาสตร์ศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

ธีระชัย นนทพิทักษ์. (2530). *องค์ประกอบบางประการที่มีอิทธิพลต่อทักษะกระบวนการทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จังหวัดอุดรธานี*. วิทยานิพนธ์ ศศ.ม. (วิทยาศาสตร์ศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์. ถ่ายเอกสาร.

✓ ธีชกร สุวรรณจรัส. (2540). *การประมาณค่าความเชื่อมั่นของแบบทดสอบ เอ็ม อี คิว วัดความสามารถในการแก้ปัญหา วิชาบัญชีกิจการบริการ ที่มีผู้ตรวจให้คะแนน 1 คน*. ปรินญาณินพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

น้อยทิพย์ ศัตตศาสตร์. (2521). *การศึกษาความสัมพันธ์ของทักษะกระบวนการทางวิทยาศาสตร์ขั้นพื้นฐานความสามารถในการแก้ปัญหาและผลสัมฤทธิ์ทางการเรียนวิทยาศาสตร์ของนักเรียนชั้นประถมศึกษาปีที่ 4*. วิทยานิพนธ์ ค.ม. (การประถมศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.

นิคม ทาแดง และสุจินต์ วิศวรธีรานนท์. (2529). *เอกสารการสอนชุดวิชาวิทยาศาสตร์ 3 หน่วยที่ 1*. พิมพ์ครั้งที่ 2. กรุงเทพฯ : นำทางการพิมพ์.

บุญชม ศรีสะอาด. (2534, มกราคม – เมษายน). "การวิเคราะห์ข้อสอบแบบอื่นนอกเหนือจากแบบเลือกตอบ," *วารสารวัดผลการศึกษา*. 12(36) : 79 – 84.

- บุญเชิด ภิญโญนนท์พงษ์. (2521). *การวัดและประเมินผลการศึกษา : ทฤษฎีและการประยุกต์*. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- _____. (2524). *ทฤษฎีการทดสอบ*. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- _____. (2525). *เอกสารประกอบการสอนวิชา วัดผล 301 การสร้างแบบทดสอบ 1*. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- _____. (2526). *การทดสอบแบบอิงเกณฑ์ : แนวคิดและปฏิบัติ*. กรุงเทพฯ : โอเดียนสโตร์.
- _____. (2533). "Congeneric part Reliability," *วารสารการวัดผลการศึกษา*, 12(34).
- _____. (2544). *การประเมินการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ : แนวคิดและวิธีการ*. กรุงเทพฯ : สำนักงานคณะกรรมการการศึกษาแห่งชาติ.
- _____. (2545). "คุณภาพของเครื่องมือวัด," ใน *ประมวลสาระชุดวิชา การพัฒนาเครื่องมือสำหรับการประเมินการศึกษา หน่วยที่ 3*. หน้า 65 -- 153. กรุงเทพฯ : มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- ปรีชา เจตินัย. (2531). *การเปรียบเทียบทักษะกระบวนการทางวิทยาศาสตร์และเจตคติทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่มีรูปแบบการจัดชั้นเรียนและแผนการเรียนแตกต่างกันในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา เขตการศึกษา 11 ปีการศึกษา 2530*. ปรินญานิพนธ์ กศ.ม. (การมัธยมศึกษา) มหาสารคาม : มหาวิทยาลัยศรีนครินทรวิโรฒ มหาสารคาม. ถ่ายเอกสาร.
- ปรีชา วงศ์ชูศิริ และคณะ. (2525). *เอกสารหน่วยการเรียนรู้การสอน : ธรรมชาติของวิทยาศาสตร์*. กรุงเทพฯ : การศาสนา.
- พรรณี เข็มมสมบูรณ์. (2543). *การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์ ที่มีจำนวนผู้ตรวจและวิธีการตรวจที่แตกต่างกัน*. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- พวงรัตน์ ทวีรัตน์. (2530). *การสร้างและพัฒนาแบบทดสอบวัดผลสัมฤทธิ์*. สำนักทดสอบทางการศึกษาและจิตวิทยา, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- พิศาล สร้อยรุห์รา. (2529). *เอกสารการสอนชุดวิทยาศาสตร์ 3 หน่วยที่ 5*. พิมพ์ครั้งที่ 2. กรุงเทพฯ : นำงการพิมพ์.
- ไพรัตน์ วงษ์นาม. (2533). *สัมประสิทธิ์การอ้างอิงสรุปของแบบทดสอบความเรียง*. วิทยานิพนธ์ ค.ด. (การวิจัยการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ไพศาล หวังพานิช. (2528). *การวัดผลการศึกษา*. กรุงเทพฯ : ไทยวัฒนาพานิช.
- ภาพ เลหาไพบูลย์. (2537). *แนวการสอนวิทยาศาสตร์*. กรุงเทพฯ : ไทยวัฒนาพานิช.
- มนวิภา อ่อนศรี. (2541). *การศึกษาความสัมพันธ์ระหว่างทักษะกระบวนการทางวิทยาศาสตร์กับความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3*. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- มณีรัตน์ เกตุไสว. (2540). *ผลการจัดกิจกรรมทางการทดลองทางการเรียนวิชาฟิสิกส์ด้านมโนคติ วิทยาศาสตร์และทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการของนักเรียน ชั้นมัธยมศึกษาปีที่ 4*. ปรินญาณิพนธ์ กศ.ม. (วิทยาศาสตร์ศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- มหาวิทยาลัย, ทบวง. คณะอนุกรรมการพัฒนาการสอนและผลิตวัสดุอุปกรณ์วิทยาศาสตร์. (2525). *ชุดการเรียนรู้ การสอนสำหรับครูวิทยาศาสตร์ เล่ม 1-3*. กรุงเทพฯ : ทบวงมหาวิทยาลัย.
- มานิช วาทะพุทกณะ. (2523). *สัมฤทธิ์ผลด้านทักษะกระบวนการทางวิทยาศาสตร์และพฤติกรรมด้าน ทักษะปฏิบัติของนักเรียนชั้นมัธยมศึกษาปีที่ 2 (ม.2) โรงเรียนบุญวัฒนา จ. นครราชสีมา*. วิทยานิพนธ์ ศศม. (การมัธยมศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์. ถ่ายเอกสาร.
- เยาวดี วิบูลย์ศรี. (2528). *หลักการวัดผลและการสร้างข้อสอบ*. กรุงเทพฯ : ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- รุจิรี ภูสาระ. (2520). *เอกสารประกอบการบรรยายวิชา ED. 337*. กรุงเทพฯ : คณะศึกษาศาสตร์ มหาวิทยาลัย ศรีนครินทรวิโรฒ.
- รวีวรรณ อังคนรักษ์พันธ์. (2531). *การสร้างแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์*. ปรินญาณิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย ศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- วรรณทิพา รอดแรงคำ และพิมพ์พันธ์ เตชะคุปต์. (2532). *ทักษะกระบวนการทางวิทยาศาสตร์สำหรับครู*. กรุงเทพฯ : เดอะมาสเตอร์กรุ๊ปแมเนจเม้นท์.
- วิญญา วิชาลาภรณ์. (2533). *การสร้างแบบทดสอบเพื่อการวิจัยและพัฒนาการเรียนการสอน*. กรุงเทพฯ : อักษราพัฒนา.
- วัฒนชัย จันทรวิบูล. (2538). *การสร้างแบบทดสอบภาคปฏิบัติวัดทักษะกระบวนการทางวิทยาศาสตร์ ชั้นพื้นฐาน*. ปรินญาณิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย ศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- วิษชุดา งามอักษร. (2541). *การศึกษาผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์ ทักษะกระบวนการทาง วิทยาศาสตร์ขั้นบูรณาการ และความสามารถในการคิดอย่างมีเหตุผลของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 โดยการสอนแบบ เอส ซี เอส กับการสอนตามคู่มือครู*. ปรินญาณิพนธ์ กศ.ม. (การมัธยมศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- ส. วาสนา ประवालพุกษ์. (2533). "การพิจารณาความเชื่อมั่นของแบบทดสอบการปฏิบัติ," *วารสารการวัดผล การศึกษา*. 12(35).
- ส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, สถาบัน. (2524). *ทักษะกระบวนการทางวิทยาศาสตร์*. กรุงเทพฯ : สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.
- _____. (2518). *รายงานการสร้างแบบทดสอบทักษะกระบวนการทางวิทยาศาสตร์*. กรุงเทพฯ : หน่วยทดสอบและประเมินผล, สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.

- ส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี, สถาบัน. (2519). *เอกสารประกอบการสอนวิชาวิทยาศาสตร์*
กายภาพ. กรุงเทพฯ : สถาบันส่งเสริมการสอนวิทยาศาสตร์และเทคโนโลยี.
- _____. (2534). *ของเล่นเชิงวิทยาศาสตร์หลากหลาย*. กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว.
- _____. (2536). *วิทยาศาสตร์กับการแก้ปัญหา*. พิมพ์ครั้งที่ 3. กรุงเทพฯ : โรงพิมพ์คุรุสภาลาดพร้าว.
- สมชัย โกมลและคณะ. (2525). *การสร้างชุดการสอนทักษะกระบวนการทางวิทยาศาสตร์สำหรับครู*.
 คณะศึกษาศาสตร์ มหาวิทยาลัยขอนแก่น.
- สมพงษ์ อิศวฤกษ์. (2542). *การเปรียบเทียบความเชื่อมั่นของแบบทดสอบที่มีวิธีการตรวจ และ*
ความยาวของแบบทดสอบแตกต่างกัน. ปรินซ์นิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ :
 บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุมาลี ดำรงไชย. (2537). *การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์และทักษะกระบวนการ*
ทางวิทยาศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ที่สอนโดยใช้แบบฝึกทักษะการทดสอบจาก
วัสดุในท้องถิ่นกับการสอนตามคู่มือครู. ปรินซ์นิพนธ์ กศ.ม. (วิทยาศาสตร์ศึกษา) กรุงเทพฯ :
 บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุวัฒน์ นิยมคำ. (2531). *ทฤษฎีและทางปฏิบัติในการสอนวิทยาศาสตร์แบบสืบเสาะหาความรู้เล่ม 1-2*.
 กรุงเทพฯ : เจเนอรัลบุ๊คส์เซนเตอร์.
- เสริม ทัดศรี. (2522). *การวิเคราะห์องค์ประกอบความสามารถทางการเรียนวิทยาศาสตร์ ของนักเรียน*
ชั้นมัธยมศึกษาปีที่ 1 (ม.1) ในจังหวัดสงขลา. ปรินซ์นิพนธ์ กศ.ด. (การมัธยมศึกษา) กรุงเทพฯ :
 บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- แสงศิริ ศิริมงคล. (2529). *การศึกษาผลสัมฤทธิ์ทางการเรียนและทักษะกระบวนการทางวิทยาศาสตร์ของ*
นักเรียนชั้นมัธยมศึกษาปีที่ 2 ที่สอนโดยทบทวนความรู้พื้นฐานด้วยบทเรียนและด้วยครู
ทบทวน. ปรินซ์นิพนธ์ กศ.ม. (การมัธยมศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย
 ศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- เสาวนีย์ เกียรติ. (2540). *คู่มือการอบรมการใช้แฟ้มสะสมผลงานนักเรียน*. กรุงเทพฯ : เนติกุลการพิมพ์.
- อนันต์ จันทร์กวี. (2525). *ผลการใช้คำถามของครูที่มีต่อทักษะกระบวนการทางวิทยาศาสตร์ ผลสัมฤทธิ์*
และทัศนคติของนักเรียน ม.ศ. 2 และ ม.2. ปรินซ์นิพนธ์ กศ.ด. (การวิจัยและพัฒนาหลักสูตร)
 กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- อุษณีย์ บัวศิริพันธุ์. (2543). *การเปรียบเทียบค่าสัมประสิทธิ์การสรุปอ้างอิงของแบบทดสอบวิชา*
คณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจ และประสบการณ์ของผู้ตรวจแตกต่างกัน.
 ปรินซ์นิพนธ์ กศ.ม. (การวัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย
 ศรีนครินทรวิโรฒ. ถ่ายเอกสาร..

- American Association for The Advancement of Science. (AAAS.) (1970). **Science A Process Approach**. New York : Comantary for Teacher. AAAS. Xerox.
- Anderson, C.R. (1978, August). "The Effectiveness of A Simulation Learning Game in Teaching Consumer Credit to Senior Approach to Instruction," **Dissertation Abstract International**. 31(2) : 127-132.
- Anderson, O.Roger. (1976). **The Experience of Science : A New Perspective for Laboratory Teaching**. New York : Teacher College Columbia University.
- Anthony, J. Nitko. (1983). **Educational assessment of student**. 2nd ed. Englewood Cliffs, N.J. : Prentice-Hall.
- Brennan, R.L. and Kane, M.T. (1979). "Generalizability Theory : A review," **New Directions for Testing and Measurement**. 4(2) : 33-51.
- Brennan, R.L. (1983). **Element of Generalizability Theory**. Iowa : The American College Testing Program.
- Burns, Joseph C. And others. (1985, February). "Development of a Integrated Process Skill Test : TIPS," **Journal of Research in Science Teaching**. 22(2) : 169-177.
- Cardinet, J., and Allal, L. (1983, September). "Estimation of generalizability parameters," **New Directions for Testing And Measurement**. 18(5) : 17-48.
- Cronbach, L.J., Rajaratnam, N. And Gleser, G.C. (1963, August). "Theory of generalizability : A Liberalization of Reliability theory," **The British Journal of Statistical Psychology**, 16(1) : 137-163.
- Cronbach, L.J. et al. (1972). **The Dependability of Behavioral Measurement : Theory of Generalizability For Score and Profiles**. New York : John Wiley & Son, Inc.
- Courtney, E.W. (1996). "Alternative Assessment Techniques at the Elementary and Secondary Levels," **เอกสารประกอบการอบรมเชิงปฏิบัติการ. คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์**.
- Cunningham, George K. (1986). **Educational and Psychological Measurement**. New York : Macmillan.
- Duren, Redeney L. (1978, January - March). "Measuring the Process of Science Objectives," **Science Education**. 62(1) : 24-25.
- Doty, Lille La Verine Carson. (1986, May). "A Study Comparing the Influence of Inquiry and Traditional Science Instruction Methods on Science Achievement, Attitudes Toward Science, and Integrated Process Skills in Ninth-Grade Students and the Relationship between Sex, Race, Past Performance in Science, Intelligence And Achievement," **Dissertation Abstracts International**. 46(11) : 3311-A.
- Feldt, L. S. and Brennan, R. L. Reliability, in Linn, R. L. (Ed.). (1989) **Educational Measurement**, New York : American Council on Educational : Macmillan Publishing Company.

- Garland, Nell A.C. and Others. (1973). **Elementary Science Learning by Investigation**. 2nd ed. New York Rand McNally.
- Garner and Engelhard, Jr. (1999, March). "Gender differences in Performance on multiple-choice and Constructed Respond Mathematics Items," **Applied Measurement in Education**. 12(1) : 29-51.
- Gay, L.R. (1992). **Educational Evaluation and Measurement Competencies for Analysis and Application**. 2nd ed. New York : Merrill.
- Good, Carter V. (1973). **Dictionary of Education**. 3rd ed. New York ; McGraw-Hill.
- Gosbi, Ali Mohamed. (1982, December). "A Study of the Understanding of Process in Relation to Piaget Cognitive Development at the Formal Level, and Other Variables Among Prospective Teachers and College Science Majors," **Dissertation Abstracts International**. 43(6) : 1914-A.
- Gronlund, Norman E. (1976). **Measurement and Evaluation in Testing**. 3rd ed. New York : Macmillan.
- _____. (1977). **Constructing Achievement Test**. 2nd ed. Prentice-Hall, Inc.
- Harkoos, Gerry D. and John E. Penick. (1983, October). "The Influence of Classroom Climate on Science Process And Content Achievement of Community College Students," **Journal of Research in Science Teaching**. 30(3) : 629-637.
- Hopkins, C. D., and Antes, R. L. (1990) **Classroom Measurement and Evaluation**. 3rd ed. Itasca, IL: F. E. Peacock.
- Hoyt, C. (1941, June). "Test reliability obtained by analysis of Variance," **Psychometrika**. 6(3) : 153 - 160.
- Ibrahim, A.M. (1984, August). "The dependability of need assessment data : An application of Generalizability Theory to the rating of the educational goals," **Dissertation Abstracts International**. 45(2) : 499-A.
- Isonio Steven. (1991, December). "Judgments of Placement Writing Samples at Golden West. College : an Education of Inter-rater Reliability," **NISC DISC REPORT**. ERIC 1985-1999.
- Jacobson, Willard J. and Bergman Abby Barry. (1980). **Science for Children**. Englewood Cliffs, New Jersey : Prentice-Hall.
- James, L.R. Demaree, R.G., Wolf, G. (1993, October). " r_{wg} : An assessment of within-group interrater Agreement," **Journal of applied psychology**. 78(2) : 306-309.
- Klein and Stecher. (1998, May). "Analytic Versus Holistic Scoring of Science Performance Tasks." **Applied Measurement in Education**. 11(2) : 131-137.
- Klopfer, L.E. (1971, Summer). "Evaluation of Learning in Science," **An Handbook on Formative and Summative Evaluation of Student Learning**. New York : McGraw-Hill Book Co.
- Kunnan, A.J. (1992, November). "An investigation of criterion-referenced test using G-theory, and factor And cluster analysis," **Language testing**. 9 (1) : 30-49.

- ✓ Lane, S. and others. (1996, Spring). "Generalizability and Validity of a Mathematics Performance Assessment," *Journal of Educational Measurement*. 33(1) : 71-91.
- Lunstrom, Donald and Lawrence Lowery. (1978). "Process pattern and Structural Themes in Science," *In Inquiry Technique for Teaching Science*. William D. Romey. Englewood Cliffs, N.J. : Prentice-Hall.
- Macready, G.B. (1983, Spring). "The use of Generalizability Theory for Assessment Relations among items Within Domains in Diagnostic Testing," *Applied Psychological Measurement*. 7(3) : 149-157.
- Marshall, John Clark and Loyde Wesley , Hales , (1971). *Classroom Test Construction*. Massachusetts : Addison-Wesley Publishing Company.
- Mehrens, William A. and Irvin J. Lehmann. (1973). *Measurement and Evaluation in Education And Psychology*. New York : Rinehart and Winston Inc.
- Molitor, Loretta. (1972, May). "an Investigation to Develop an Instrument to Evaluate the Science Skills of Inference Prediction and Verification of Children in Grade Four, Five and Six," *Dissertation Abstracts International*. 32(11) : 6241-A-6242-A.
- O'Brien, R.M. and Jones, B. (1986, October). "The Reliability of School – Aggregate Variables : An Application of Generalizability Theory," *Journal of Research and Development in Education*. 20(2) : 21-27.
- Okey, James R. And Ronald L. Field. (1973). *Basic Process Skills Program*. Bloomington : Indiana University.
- Padilla, Michael J. and Others. (1983, March). "The Relationship Between Science Process Skills and Formal Thinking," *Journal of Research in Science Teaching*. 20(3) : 239-246.
- Pappelis, C.K. and Others. (1980, March). "Can Instruction Improve Science Process Skill of Premedical and Pre dental Students," *Journal of Research in Science Teaching*. 17(1) : 25-29.
- Perez, Carolina V. (1978, December). "The Development and Evaluation of a Test of Science Process for Use in The Philippines," *Dissertation Abstracts International*. 39(6) : 3496-A.
- Peterson, K.D. (1978, March). "Scientific Inquiry Training for high School Student," *Journal of Research In Science Teaching*. 15(3) : 153.
- Petus, A.M. and C.D. Haley. (1980, April). "Identifying Factors Related to Science Process Skill Performance Levels," *School Science and Mathematics*. 30(2) : 273-276.
- Pomplun and Sundbye. (1999, September). "Gender Differences in Constructed Response Reading items," *Applied Measurement in Education*. 12(1) : 95-109.
- Powers and Fowles. (1998, Spring). "Effects of Pre-examination Disclosure of Essay Topics," *Applied Measurement in Education*. 11(2) : 139-157.

Robinson, Richard Wayne. (1973, September). "The Development of Item Which Access the Process of Controlling Variables and Interpreting Data," *Dissertation Abstracts International*. 35(3) : 1522-A.

Roth, Wolff Michael and Anita Roychoudhury. (1993, February). "The Development of Science Process Skills in Authentic Contexts," *Journal of Research in Science Teaching*. 30(2) : 127-152.

Rubin, Rochelle Leventhal. (1989, April). "Using a Systematic Modeling Teaching Strategy to Promote the Development of Integrated Science Process Skills and Formal Cognitive Reasoning Ability (Reasoning)," *Dissertation Abstracts International*. 50(11) : 3469-A.

Scharmann, Lawrence C. (1989, November). "Developmental Influences of Science Process Skill Instruction," *Journal of Research in Science Teaching*. 26() : 715-726.

Showalter, V.M. (1974, March). "What Is Unified Science Education 7 (Part 6) : Programs objectives And Scientific Literacy," *Prism II*, 2(4) : 1-8.

Smith, P.L. (1978, April). "Sampling Error of Variance Components in Small Sampling Multifacet Generalizability Studies," *Journal of Educational Statistics*. 3(4) : 319-346.

✍ Spearman, C. (1910, March). "Coefficient of Correlation Calculated from Faulty Data," *British Journal of Psychology*, 3(2) : 271 – 295.

Stenmark, J.K. (1991). *Mathematics Assessment : myths, models, good question, and Practical Suggestion*. University of California : The Nation Council of Teachers of Mathematics, INC.

Strawitz, B.M. and M.R. Malone (1987, January). "Preservice Teacher's Acquisition and Retention of Integrated Science Process Skills : A Comparison of Teacher-Directed and Self-Instructional Strategies," *Journal of Research in Science Teaching*. 24() : 53-60.

Sund, Robert B. and A.J. Picard. (1972). *Behavioral Objectives and Evaluational Measure : Science And Mathematics*. Ohio, Charlee E. Merrill.

Sund, Robert B. and leslie W. Trowbridge. (1967). *Teaching Science by Inquiry in the Secondary School*. Columbus, Ohio:Charles E.Meerrill Co.

Tasse, Mare J. And others. (1994, February). "Test-Retest and Inter-rater Reliability of the Quebec Adaptive Behavior Scale." *NISC DISC REPORT*. ERIC 1985-1999.

Ward, Annie W. (1999). *Assessment in Classroom*. p. 421. Wadsworth Publishing Company.

Wiersma and Jure. (1985). *Education Measurement and Testing*. Allyn and Bacon, Inc.

✍ Woodruff, D.J. & Feldt, L.S. (1986, September). "Test for equality of several alpha coefficients when their sample estimates are dependent." *Psychometrika*. 51(2) : 393-413.

ภาคผนวก

ภาคผนวก

1. รายนามผู้เชี่ยวชาญตรวจสอบคุณภาพของเครื่องมือ
2. ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญ (IOC)
3. ค่าความยาก (p) ค่าอำนาจจำแนก (t) ของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์
ชั้นบูรณาการ
4. คู่มือการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
5. ผลการตรวจให้คะแนนแบบทดสอบ

ตาราง 19 ดัชนีความสอดคล้องของการประเมินจากผู้เชี่ยวชาญในการตรวจสอบคุณภาพของเครื่องมือ (IOC)

ทักษะ	ข้อ	ผู้เชี่ยวชาญคนที่					รวม	IOC	แปลผล
		1	2	3	4	5			
การตั้งสมมติฐาน	1.1	1	1	0	1	0	3	0.6	ใช้ได้
	1.2	0	1	1	1	1	4	0.8	ใช้ได้
	1.3	1	1	1	1	1	5	1.0	ใช้ได้
การกำหนดนิยาม เชิงปฏิบัติการ	2.1	1	1	1	1	1	5	1.0	ใช้ได้
	2.2	1	1	1	1	0	4	0.8	ใช้ได้
	2.3	1	1	1	1	1	5	1.0	ใช้ได้
	2.4	1	1	1	1	1	5	1.0	ใช้ได้
การกำหนดและควบคุม ตัวแปร	3.1	1	1	1	0	1	4	0.8	ใช้ได้
	3.2	0	1	1	1	0	3	0.6	ใช้ได้
	3.3	1	1	1	0	1	4	0.8	ใช้ได้
	3.4	1	1	1	0	1	4	0.8	ใช้ได้
	3.5	1	1	1	1	0	4	0.8	ใช้ได้
การทดลอง	4.1	0	1	0	1	1	3	0.6	ใช้ได้
	4.2	0	1	1	0	1	3	0.6	ใช้ได้
	4.3	1	1	0	1	1	4	0.8	ใช้ได้
การตีความหมายข้อมูล และลงข้อสรุป	5.1	0	1	1	1	1	4	0.8	ใช้ได้
	5.2	1	1	1	1	0	4	0.8	ใช้ได้
	5.3	0	1	1	1	1	4	0.8	ใช้ได้
	5.4	1	1	0	1	1	4	0.8	ใช้ได้

ตาราง 20 ค่าความยาก (p) ค่าอำนาจจำแนก (r_{xy}) ของแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์
ชั้นบูรณาการ

ทักษะ	ข้อ	ค่าความยาก (p)	ค่าอำนาจจำแนก (r_{xy})	ผลการคัดเลือก
การตั้งสมมติฐาน	1.1	0.316	.668**	ใช้ได้
	1.2	0.368	.636**	ใช้ได้
	1.3	0.421	.684**	ใช้ได้
การกำหนดนิยาม เชิงปฏิบัติการ	2.1	0.395	.620**	ใช้ได้
	2.2	0.474	.645**	ใช้ได้
	2.3	0.368	.332	คัดออก
	2.4	0.553	.750**	ใช้ได้
การกำหนดและควบคุม ตัวแปร	3.1	0.474	.671**	ใช้ได้
	3.2	0.316	.678**	ใช้ได้
	3.3	0.509	.706**	ใช้ได้
	3.4	0.561	.845**	ใช้ได้
	3.5	0.246	.476*	คัดออก
การทดลอง	4.1	0.322	.582**	ใช้ได้
	4.2	0.401	.391	คัดออก
	4.3	0.401	.624**	ใช้ได้
การตีความหมายข้อมูล และลงข้อสรุป	5.1	0.447	.649**	ใช้ได้
	5.2	0.789	.677**	ใช้ได้
	5.3	0.421	.529*	ใช้ได้
	5.4	0.474	.781**	ใช้ได้

หมายเหตุ ทักษะการทดลองเป็นทักษะที่ต้องใช้เวลามาก หลังจากที่ย่านวนเวลาดูแล้ว ถ้าในแบบทดสอบใช้ทักษะนี้มากกว่า 1 ข้อ จะใช้เวลาเกินกว่าที่กำหนด ดังนั้น ทักษะนี้ผู้วิจัยจึงเลือกเอาเพียงข้อเดียว คือ ข้อ 4.3 ที่สามารถจำแนกกลุ่มสูงและกลุ่มต่ำได้สูงกว่าข้อ 4.1

คู่มือการตรวจให้คะแนน แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ

คำชี้แจง

1. คู่มือการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ ประกอบด้วย 2 องค์ประกอบด้วยกัน คือ
 - 1.1 แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
 - 1.2 แนวทางในการตอบและกฎเกณฑ์การให้คะแนน (Scoring Rubric)
2. การตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ กำหนดให้ตรวจความถูกต้องของคำตอบในแต่ละข้อโดยใช้แนวทางในการตอบประกอบการพิจารณาให้คะแนน และนำผลการตรวจ มาจัดระดับคุณภาพของผู้เข้าสอบแยกเป็นรายทักษะโดยใช้กฎเกณฑ์การให้คะแนน (Scoring Rubric) ประกอบการพิจารณา กำหนดระดับคุณภาพเป็น 4 ระดับ ทุกทักษะกระบวนการย่อย

1.1 แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการ

คำชี้แจง

1. แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการฉบับนี้ มีคำถามทั้งหมด 15 ข้อ ใช้เวลาในการทำแบบทดสอบ 120 นาที ห้ามขีดเขียนข้อความใด ๆ ลงในแบบทดสอบโดยเด็ดขาด
2. แบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ขั้นบูรณาการฉบับนี้ ประกอบด้วยทักษะกระบวนการย่อยจำนวน 5 ทักษะ คือ ทักษะการตั้งสมมติฐาน ทักษะการกำหนดนิยามเชิงปฏิบัติการ ทักษะการกำหนดและควบคุมตัวแปร ทักษะการทดลอง และทักษะการตีความหมายข้อมูลและลงข้อสรุป
3. แบบทดสอบทุกข้อเป็นแบบอัตนัย โดยที่แต่ละข้อกำหนดสถานการณ์ให้นักเรียนศึกษาแล้วจึงเขียนตอบลงในกระดาษคำตอบที่ได้จัดเตรียมไว้ในแต่ละข้อ ดังตัวอย่าง ข้อ (0)

ตัวอย่างแบบทดสอบ

ข้อ (0) สมมติว่านายวิจิตรเคยรับประทานปูนามาแล้ว 6 ครั้ง ผลปรากฏว่า หลังจากรับประทานปูนาแล้วทุกครั้งเขาจะมีอาการท้องเดิน ถ้าต้องการทดลองว่านายวิจิตรรับประทานกุ้งแล้วจะเกิดผลอย่างไร สมมติฐานของการทดลองในครั้งนี้ควรตั้งอย่างไร

.....

ฯลฯ

4. ควรศึกษาข้อความในแต่ละสถานการณ์ให้เข้าใจเสียก่อนแล้วจึงลงมือตอบ
5. ถ้าพบข้อยากอย่างท้อใจ เว้นข้ามไปทำข้ออื่น ๆ ต่อไปก่อนเมื่อมีเวลาเหลือจึงย้อนกลับมาทำใหม่
6. อย่าเปิดแบบทดสอบจนกว่าจะได้รับสัญญาณให้ลงมือทำ
7. เมื่อได้รับสัญญาณเตือนหมดเวลาในการทำข้อสอบ ให้นักเรียนหยุดการทำข้อสอบในทันที

ข้อ 1 ทักษะการตั้งสมมติฐาน

1.1 จากการทดลองละลายสาร A ในน้ำเย็นปรากฏว่า สาร A 21 กรัม ละลายได้พอดีในน้ำ 250 ลูกบาศก์เซนติเมตร จากการทดลองนี้ ถ้าผู้ทดลองมีความเชื่อว่า การอุ่นสารละลายให้ร้อนขึ้นจะช่วยให้สาร A ละลายได้เพิ่มขึ้นอีก ผู้ทดลองควรตั้งสมมติฐานว่าอย่างไร

.....

.....

.....

1.2 นำลูกกลมตันที่ทำด้วยวัตถุต่างชนิดกัน ได้แก่ วัตถุ A , B และวัตถุ C ซึ่งมีเส้นผ่าศูนย์กลางเท่ากันมา 3 ลูก ปล่อยให้กลิ้งบนพื้นเอียงเดียวกัน แล้วจับเวลาที่ลูกกลมแต่ละชนิดใช้ในการกลิ้งจากปลายบนสุดถึงปลายล่างสุดของพื้นเอียง ทำซ้ำแบบเดิม 3 ครั้ง หาค่าเฉลี่ยและนำมาเปรียบเทียบกัน กรณีดังกล่าว ถ้านักเรียนเป็นผู้ทดลองจะตั้งสมมติฐานอย่างไร

.....

.....

.....

1.3 จงศึกษาข้อมูลจากตารางต่อไปนี้

ชนิดของสาร	ความสามารถในการละลายของสารที่อุณหภูมิห้อง (กรัม)	
	ในน้ำ 100 ลบ.ซม.	ในเอทิลแอลกอฮอล์ 100 ลบ.ซม.
A	140	5
B	39	1
C	200	0.6
D	36	0
E	6	0

ข้อมูลในการทดลองจากตารางนี้ ควรตั้งสมมติฐานอย่างไร

.....

.....

.....

ข้อ 2 ทักษะการกำหนดนิยามเชิงปฏิบัติการ

2.1 คำว่า "น้ำใส" จะกำหนดนิยามอย่างไร ให้สามารถใช้เป็นแนวในการสังเกตและวัดได้ตรงกัน เมื่อต้องการจำแนกน้ำใสออกจากน้ำที่มีลักษณะอื่น ๆ

.....

.....

.....

2.2 ในการศึกษาเรื่อง นก จากประโยคที่ว่า “นกเป็นสัตว์ปีก ส่วนมากบินได้” คำว่า “บินได้” ของนก ควรนิยามอย่างไร จึงจะสังเกตและวัดได้ตรงกัน

.....

.....

.....

2.3 คำว่า “ก๊าซคาร์บอนไดออกไซด์” ควรนิยามอย่างไร จึงจะสังเกตและวัดได้ตรงกัน เมื่อต้องการศึกษาปฏิกิริยาของก๊าซคาร์บอนไดออกไซด์

.....

.....

.....

ข้อ 3 ทักษะการกำหนดและควบคุมตัวแปร

3.1 ชนงูได้ทำการทดลองโดยใช้คลอรีนคลอไรด์ 1 ช้อนเบอร์ 1 เท่ากัน ลงในบีกเกอร์ใบเล็กขนาดเดียวกัน 4 ใบ ซึ่งแต่ละใบบรรจุน้ำที่มีปริมาณแตกต่างกันและคนให้ทั่ว หลังจากที่ใช้คลอรีนคลอไรด์ละลายหมด เขาก็วัดอุณหภูมิของน้ำภายในบีกเกอร์แต่ละใบ จากข้อความดังกล่าวจงระบุ ตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุม

ตัวแปรต้น

ตัวแปรตาม

ตัวแปรที่ต้องควบคุม

.....

.....

.....

.....

3.2 เพื่อศึกษาว่าปริมาณของวิตามินเอที่หนูได้รับมีผลต่อน้ำหนักของหนูหรือไม่ นักวิทยาศาสตร์ได้ทดลองเลี้ยงหนู โดยแบ่งหนูออกเป็น 5 กลุ่ม แต่ละกลุ่มจะได้รับอาหารที่เหมือนกัน แต่ได้รับปริมาณของวิตามินที่แตกต่างกัน หลังจากเลี้ยงหนูไว้ 3 สัปดาห์ เขาก็ชั่งน้ำหนักของหนูแต่ละตัว จากข้อความดังกล่าวจงระบุ ตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุม

ตัวแปรต้น

ตัวแปรตาม

ตัวแปรที่ต้องควบคุม

.....

.....

.....

.....

3.3 แม่เหล็กไฟฟ้าจะดูดจำนวนตะปูได้มากขึ้นหรือไม่ ถ้าแม่เหล็กไฟฟ้านั้นมีจำนวนแบตเตอรี่เพิ่มขึ้น จากข้อความดังกล่าวจงระบุ ตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุม

ตัวแปรต้น

ตัวแปรตาม

ตัวแปรที่ต้องควบคุม

.....

.....

.....

.....

3.4 การงอกของเมล็ดข้าวโพดในเวลาที่แตกต่างกันขึ้นอยู่กับปริมาณของน้ำที่เมล็ดข้าวโพดได้รับ จากข้อความดังกล่าวจงระบุ ตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุม

ตัวแปรต้น

ตัวแปรตาม

ตัวแปรที่ต้องควบคุม

.....

.....

.....

.....

ข้อ 4 ทักษะการทดลอง

ชาวสวนคนหนึ่งมีที่ดินอยู่หลายแปลง เขาสังเกตว่าที่ดินแต่ละแปลงที่ใช้ในการปลูกพืช ต้องรดน้ำในปริมาณที่แตกต่างกัน เขาเกิดความสงสัยว่า ดินต่างชนิดกันจะอุ้มน้ำไว้ในปริมาณที่ต่างกันหรือไม่ จึงได้ตั้งสมมติฐานว่า “ชนิดของดินมีผลต่อปริมาณของน้ำที่อุ้มน้ำไว้”

4.1 ถ้านักเรียนเป็นชาวสวนคนนั้นจะมีวิธีออกแบบการทดลองอย่างไร

.....

.....

.....

4.2 จะใช้วัสดุอุปกรณ์ประกอบการทดลองอะไรบ้าง

.....

.....

.....

4.3 รายละเอียดขั้นตอนการดำเนินการทดลองเป็นอย่างไร

.....

.....

.....

4.4 การบันทึกผลการทดลองควรออกแบบอย่างไร

.....

.....

.....

ข้อ 5 ทักษะการตีความหมายข้อมูลและลงข้อสรุป

5.1 จงศึกษาข้อมูลในตารางที่แสดงปริมาณการจับปลาทะเลของเรือประมงลำหนึ่ง โดยใช้เครื่องมือ
อวนลากในเวลา 1 ชั่วโมง ระหว่างปี พ.ศ. 2508 – 2520 แล้วตอบคำถามในข้อ 5.1.1 – 5.1.2

พ.ศ.	จำนวนปลาที่เรือจับได้ (kg)
2508	232
2509	132
2510	115
2511	106
2512	103
2513	97
2514	66
2515	63
2516	52
2517	58
2518	47
2519	57
2520	47

5.1.1 ตั้งแต่ปี พ.ศ. 2508 – 2520 จำนวนปลาที่จับได้ในแต่ละชั่วโมงเปลี่ยนแปลงอย่างไร

.....

.....

.....

5.1.2 จากข้อมูลข้างต้นลงข้อสรุปได้ว่อย่างไร

.....

.....

.....

5.2 จากการทดลองละลายสาร A ในของเหลว B จำนวน 50 ลูกบาศก์เซนติเมตร ที่อุณหภูมิต่าง ๆ กัน ได้ผลการทดลองดังตาราง ให้ใช้เป็นข้อมูลตอบคำถามข้อ 5.2.1 – 5.2.2

อุณหภูมิของของเหลว B ($^{\circ}\text{C}$)	ปริมาณของสาร A ที่ละลายในของเหลว B (g)
20	5
30	10
40	20
50	40

5.2.1 ที่อุณหภูมิ 40°C และ 50°C สาร A ละลายในของเหลว B ได้แตกต่างกันเท่าใด

.....

.....

.....

5.2.2 จากข้อมูลในตารางเมื่ออุณหภูมิสูงขึ้น การละลายของสาร A เป็นอย่างไร

.....

.....

.....

5.3 จากการวัดความดันบรรยากาศและความหนาแน่นของอากาศที่ระดับความสูงต่าง ๆ ได้ผลการทดลองดังนี้

ความสูงจากระดับน้ำทะเล (km)	ความดันบรรยากาศ (mm ของปรอท)	ความหนาแน่นของอากาศ (g/cm^3)
0	760	0.00130
2	600	0.00100
4	470	0.00082
6	360	0.00066
8	280	0.00052
10	210	0.00041
12	160	0.00031

5.3.1 ความดันบรรยากาศที่ระดับความสูง 6 กิโลเมตร จากระดับน้ำทะเลมีค่าเท่าใด

.....

.....

.....

5.3.2 จากตารางข้อมูลข้างต้น สรุปความสัมพันธ์ระหว่างระดับความสูง ความดันบรรยากาศและความหนาแน่นของอากาศได้อย่างไร

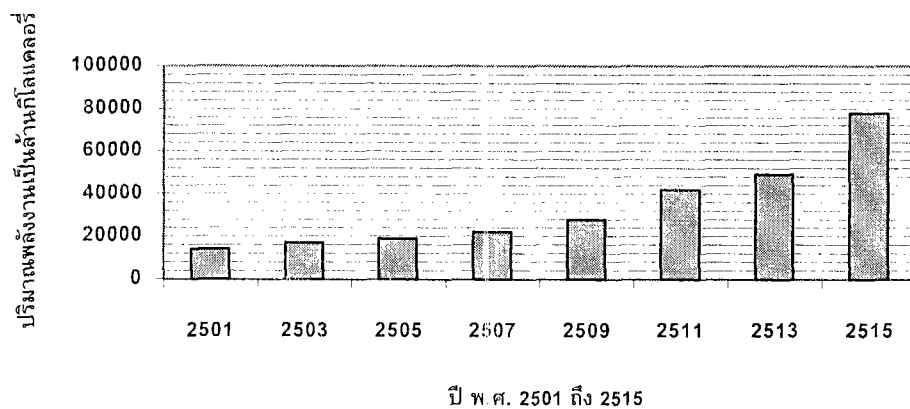
.....

.....

.....

5.4 จากการสำรวจพลังงานที่ใช้ในประเทศไทยจาก ปี พ.ศ. 2501 ถึง พ.ศ. 2515 ได้ผลดังแผนภูมิข้างล่าง ให้ใช้เป็นข้อมูลตอบคำถามข้อ 5.4.1 – 5.4.2

กราฟแสดงปริมาณพลังงานที่ใช้ในประเทศไทยจากปี พ.ศ. 2501 ถึง 2515



5.4.1 พลังงานที่ใช้ในปี พ.ศ. 2505 กับปี พ.ศ. 2515 ต่างกันเท่าใด

.....

.....

.....

5.4.2 จากข้อมูลในแผนภูมิข้างต้น สรุปความสัมพันธ์ระหว่างปริมาณพลังงานที่ใช้ในประเทศไทยกับระยะเวลาในทุก ๆ 2 ปีได้อย่างไร

.....

.....

.....

1.2 แนวทางในการตอบและกฎเกณฑ์การให้คะแนน

การตรวจให้คะแนน ข้อ 1 ทักษะการตั้งสมมุติฐาน

1. ในแต่ละข้อย่อย เช่น ข้อ 1.1, 1.2 และ 1.3 กำหนดน้ำหนักคะแนนข้อละ 1 คะแนน ถ้าตอบถูกต้องตามแนวทางในการตอบ หรือ มีส่วนใกล้เคียง หรือ มีแนวทางที่จะนำไปสู่คำตอบ ให้ข้อย่อยละ 1 คะแนน ถ้าไม่เป็นไปตามเงื่อนไขที่กำหนด ให้ 0 คะแนน

2. นำผลการตรวจให้คะแนนมาจัดระดับคุณภาพ โดยใช้กฎเกณฑ์การให้คะแนนทักษะการตั้งสมมุติฐานประกอบการพิจารณาพฤติกรรมที่บ่งชี้ความสามารถของผู้เข้าสอบ กำหนดคุณภาพเป็น 4 ระดับ บันทึกผลการให้คะแนนในภาพรวมของทักษะ

แนวทางในการตอบ

1.1 การเพิ่มอุณหภูมิให้แก่น้ำจะมีผลต่อปริมาณการละลายของสาร A

1.2 ชนิดของวัตถุที่ใช้ทำลูกกลมนั้นมีผลต่อเวลาในการกลิ้งของลูกกลมนบนพื้นเอียง

1.3 ความสามารถในการละลายของสารแต่ละชนิดขึ้นอยู่กับชนิดของตัวทำละลาย หรือ ชนิดของตัวทำละลายมีผลต่อความสามารถในการละลายของสารแต่ละชนิด

กฎเกณฑ์การให้คะแนนทักษะการตั้งสมมุติฐาน

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดชัดเจนถูกต้องทั้งหมด
3 ดี	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดชัดเจนถูกต้อง 2 ข้อใน 3 ข้อ
2 ใช้ได้	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองได้สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดชัดเจนถูกต้อง 1 ข้อใน 3 ข้อ
1 ใช้ไม่ได้	หมายถึง สามารถคาดคะเนคำตอบก่อนการทดลองไม่สอดคล้องกับตัวแปรต้นและตัวแปรตาม การแสดงเหตุผลและแนวคิดมีแนวโน้มที่จะได้คำตอบ
0	ไม่ตอบหรือตอบไม่ถูกเลย

การตรวจให้คะแนน ข้อ 2 ทักษะการกำหนดนิยามเชิงปฏิบัติการ

1. ในแต่ละข้อย่อย เช่น ข้อ 2.1, 2.2 และ 2.3 กำหนดน้ำหนักคะแนนข้อละ 1 คะแนน ถ้าตอบถูกต้องตามแนวทางในการตอบ หรือ มีส่วนใกล้เคียง หรือ มีแนวทางที่จะนำไปสู่คำตอบ ให้ข้อย่อยละ 1 คะแนน ถ้าไม่เป็นไปตามเงื่อนไขที่กำหนด ให้ 0 คะแนน
2. นำผลการตรวจให้คะแนนมาจัดระดับคุณภาพ โดยใช้กฎเกณฑ์การให้คะแนนทักษะการกำหนดนิยามเชิงปฏิบัติการประกอบการพิจารณาพฤติกรรมที่บ่งชี้ความสามารถของผู้เข้าสอบ กำหนดคุณภาพเป็น 4 ระดับ บันทึกผลการให้คะแนนในภาพรวมของทักษะ

แนวทางในการตอบ

- 2.1 น้ำใส หมายถึง น้ำที่สามารถมองเห็นวัตถุที่อยู่อีกด้านหนึ่งได้ชัดเจนเสมือนหนึ่งเห็นวัตถุ นั้นโดยตรง
- 2.2 บินได้ หมายถึง การที่ร่างกายทุก ๆ ส่วนของนกลอยอยู่ในอากาศหรือเคลื่อนที่ไปมาในอากาศ ได้ด้วยการใช้ปีกช่วยในการเคลื่อนที่
- 2.3 ก๊าซคาร์บอนไดออกไซด์ หมายถึง ก๊าซชนิดหนึ่งที่มีอยู่ในอากาศ เมื่อผ่านลงในน้ำปูนใสแล้วจะทำให้ น้ำปูนใสขุ่น

กฎเกณฑ์การให้คะแนนทักษะการกำหนดนิยามเชิงปฏิบัติการ

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง ระบุความหมายและขอบเขตของคำหรือตัวแปรได้ชัดเจน การแสดงเหตุผลถูกต้องทั้งหมด พร้อมทั้งสามารถเสนอแนะแนวทางในการตรวจสอบและวัดได้
3 ดี	หมายถึง ระบุความหมายและขอบเขตของคำหรือตัวแปรได้ชัดเจน การแสดงเหตุผลถูกต้อง 2 ข้อใน 3 ข้อ พร้อมทั้งเสนอแนะแนวทางในการตรวจสอบและวัดได้
2 ใช้ได้	หมายถึง ระบุความหมายและขอบเขตของคำหรือตัวแปรได้ชัดเจน การแสดงเหตุผลถูกต้อง 1 ข้อใน 3 ข้อ พร้อมทั้งเสนอแนะแนวทางในการตรวจสอบและวัดได้
1 ใช้ไม่ได้	หมายถึง ระบุความหมายและขอบเขตของคำหรือตัวแปรได้ชัดเจน การแสดงเหตุผลถูกต้อง 1 ข้อใน 3 ข้อ แต่ไม่สามารถเสนอแนะแนวทางในการตรวจสอบและวัดได้
0	ไม่ตอบหรือตอบไม่ถูกเลย

การตรวจให้คะแนน ข้อ 3 ทักษะการกำหนดและควบคุมตัวแปร

1. ในแต่ละข้อย่อย เช่น ข้อ 3.1 – 3.4 กำหนดน้ำหนักคะแนนข้อละ 1 คะแนน ถ้าตอบถูกต้องตามแนวทางในการตอบ หรือ มีส่วนใกล้เคียง หรือ มีแนวทางที่จะนำไปสู่คำตอบ โดยที่ผู้เข้าสอบจะต้องตอบถูกต้องให้ครบทั้งสามตัวแปรในแต่ละข้อย่อย จึงจะได้ 1 คะแนน ในข้อย่อยนั้น ถ้าไม่เป็นไปตามเงื่อนไขที่กำหนด ให้ 0 คะแนน

2. นำผลการตรวจให้คะแนนมาจัดระดับคุณภาพ โดยใช้กฎเกณฑ์การให้คะแนนทักษะการกำหนดและควบคุมตัวแปรประกอบการพิจารณาพฤติกรรมที่บ่งชี้ความสามารถของผู้เข้าสอบ กำหนดคุณภาพเป็น 4 ระดับ บันทึกผลการให้คะแนนในภาพรวมของทักษะ

แนวทางในการตอบ

3.1 ตัวแปรต้น คือ ความเข้มข้นของสารละลายคลอไรด์

ตัวแปรตาม คือ อุณหภูมิของสารละลายคลอไรด์

ตัวแปรที่ต้องควบคุม คือ ปริมาณของคลอไรด์ ,ขนาดของช้อนตักสาร , เทอร์มอมิเตอร์ที่ใช้วัดอุณหภูมิ ฯลฯ

3.2 ตัวแปรต้น คือ ปริมาณของวิตามิน เอ

ตัวแปรตาม คือ น้ำหนักของหนู

ตัวแปรที่ต้องควบคุม คือ อาหารที่ใช้เลี้ยงหนู ,ชนิดและขนาดของหนู , ระยะเวลาในการเลี้ยง, สถานที่ ฯลฯ

3.3 ตัวแปรต้น คือ จำนวนแบตเตอรี่

ตัวแปรตาม คือ จำนวนตะปูที่ถูกดูด

ตัวแปรที่ต้องควบคุม คือ ชนิดของแม่เหล็กไฟฟ้า ,ชนิดของแบตเตอรี่ ฯลฯ

3.4 ตัวแปรต้น คือ ปริมาณของน้ำ

ตัวแปรตาม คือ เวลาในการงอกของเมล็ดข้าวโพด

ตัวแปรที่ต้องควบคุม คือ ชนิดและความสมบูรณ์ของเมล็ดข้าวโพด ,ชนิดและปริมาณของดินที่ใช้ปลูกข้าวโพด ,แสงสว่าง ,อุณหภูมิ ฯลฯ

กฎเกณฑ์การให้คะแนนทักษะการกำหนดและควบคุมตัวแปร

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง กำหนดตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุมได้ถูกต้องทั้งหมด
3 ดี	หมายถึง กำหนดตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุมได้ถูกต้อง 3 ข้อใน 4 ข้อ
2 ใช้ได้	หมายถึง กำหนดตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุมได้ถูกต้อง 2 ข้อใน 4 ข้อ
1 ใช้ไม่ได้	หมายถึง กำหนดตัวแปรต้น ตัวแปรตาม และตัวแปรที่ต้องควบคุมได้ถูกต้อง 1 ข้อใน 4 ข้อ
0	ไม่ตอบหรือตอบไม่ถูกเลย

การตรวจให้คะแนน ข้อ 4 ทักษะการทดลอง

ในข้อนี้มีคุณลักษณะพิเศษคือ ข้อย่อยแต่ละข้อมีความสัมพันธ์เกี่ยวข้องกับสถานการณ์ในข้อใหญ่ จนไม่สามารถแยกคะแนนให้เป็นรายข้อย่อยเหมือนข้ออื่น ๆ ได้ อีกทั้งการตอบของผู้เข้าสอบอาจไม่ตรงกับแนวทางในการตอบที่ผู้วิจัยเสนอไว้ ผู้ตรวจต้องใช้ประสบการณ์การสอนวิทยาศาสตร์ร่วมกับกฎเกณฑ์การให้คะแนนทักษะการทดลอง ประกอบการพิจารณาพฤติกรรมการป้องกันขีดความสามารถของผู้เข้าสอบ บันทึกผลการให้คะแนนในภาพรวมของทักษะ (สามารถใช้คะแนนระดับคุณภาพเป็นคะแนนดิบ โดยไม่ต้องไปพิจารณาข้อย่อยเหมือนทักษะอื่น ๆ)

แนวทางในการตอบ

4.1 การออกแบบการทดลอง โดยการนำเอาตัวอย่างดินต่างชนิดกันมาทดสอบความสามารถในการอุ้มน้ำ การทดลองนี้จะต้องคำนึงถึงปริมาณของดิน , ปริมาณของน้ำ , ขนาดของภาชนะที่บรรจุ , ระยะเวลาในการทดลอง ฯลฯ เพื่อลดข้อผิดพลาดในการทดลอง

4.2 วัสดุอุปกรณ์ประกอบการทดลอง ได้แก่

1. กระป๋องนม หรืออาจเลือกใช้ภาชนะอื่นที่สามารถนำมาทดลองได้ ที่มีรูปร่างและขนาดเดียวกัน 6 ใบ โดยที่ 3 ใบ จะรูที่กันกระป๋องพอประมาณ
2. ดินเหนียว ดินร่วน และดินทราย
3. น้ำ

4. กระบอกตวงขนาด 100 ลูกบาศก์เซนติเมตร
5. ผ้าขาวบางขนาด 15x15 ตารางเซนติเมตร จำนวน 3 ผืน
6. นาฬิกาจับเวลา

4.3 รายละเอียดขั้นตอนการดำเนินการทดลองเป็นดังนี้

1. นำดินแต่ละชนิด คือ ดินเหนียว ดินร่วน และดินทราย มาใส่กระป๋องนมที่เจาะรูแต่ละใบ ในปริมาณ 2/3 ของกระป๋อง
2. นำกระป๋องที่ใส่ดินเหล่านี้วางซ้อนบนกระป๋องที่ไม่ได้เจาะรูแต่ละใบโดยมีผ้าขาวบางกั้นไว้ระหว่างกระป๋องทั้งสอง
3. ตวงน้ำปริมาตร 100 ลูกบาศก์เซนติเมตร แล้วเทลงในกระป๋องนมที่ใส่ดินแต่ละชนิด
4. จับเวลาของน้ำที่ไหลจากกระป๋องบนสู่กระป๋องด้านล่าง จนครบทั้งหมด
5. ปลอยให้น้ำไหลออกจากกระป๋องบนสู่โบล้างประมาณ 2 นาที แล้วนำน้ำในกระป๋องโบล้างแต่ละใบไปตวงหาปริมาตร จนครบทั้งหมด แล้วบันทึกผล
6. ทำการทดลองซ้ำตามข้อ 1-5 อีก 2 ครั้ง บันทึกผลการทดลอง

4.4 ถ้าตอบตามแนวทางข้อ 4.1-4.3 การออกแบบการบันทึกผลการทดลองควรเป็นดังนี้

ตารางบันทึกผลการทดลอง

ชนิดของดินในกระป๋อง	ปริมาตรของน้ำก่อนเท (cm ³)	ปริมาตรน้ำที่ไหลผ่านดิน (cm ³)				ปริมาตรน้ำที่เหลือในกระป๋องดิน (cm ³)			
		ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3	เฉลี่ย	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3	เฉลี่ย
ดินเหนียว	100								
ดินร่วน	100								
ดินทราย	100								

หมายเหตุ

ถ้าผู้เข้าสอบตอบหรือเสนอวิธีการทดลองที่ไม่ตรงกับแนวทางในการตอบ ให้ผู้ตรวจใช้ประสบการณ์ในการสอนวิทยาศาสตร์พิจารณาให้คะแนนได้ตามความเหมาะสม

กฎเกณฑ์การให้คะแนนทักษะการทดลอง

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง ออกแบบการทดลองได้ถูกต้องและสอดคล้องกับสมมติฐาน ระบุวัสดุอุปกรณ์ที่ใช้ในการทดลองได้ครอบคลุม อธิบายขั้นตอนการทดลองได้ถูกต้อง และเสนอแนะการบันทึกผลการทดลองได้สอดคล้องกับข้อมูล
3 ดี	หมายถึง ออกแบบการทดลองได้ถูกต้องและสอดคล้องกับสมมติฐาน ระบุวัสดุอุปกรณ์ที่ใช้ในการทดลองไม่ครอบคลุม อธิบายขั้นตอนการทดลองได้ถูกต้อง และเสนอแนะการบันทึกผลการทดลองได้สอดคล้องกับข้อมูล
2 ใช้ได้	หมายถึง ออกแบบการทดลองได้ถูกต้องและสอดคล้องกับสมมติฐาน ระบุวัสดุอุปกรณ์ที่ใช้ในการทดลองไม่ครอบคลุม แต่ไม่อธิบายขั้นตอนการทดลองหรือเสนอแนะการบันทึกผลการทดลองอย่างใดอย่างหนึ่ง
1 ใช้ไม่ได้	หมายถึง ออกแบบการทดลองไม่สอดคล้องกับสมมติฐาน ระบุวัสดุอุปกรณ์ที่ใช้ในการทดลองไม่ครอบคลุม ไม่อธิบายขั้นตอนการทดลอง และไม่เสนอแนะการบันทึกผลการทดลอง
0	ไม่ตอบหรือตอบไม่ถูกเลย

การตรวจให้คะแนน ข้อ 5 ทักษะการตีความหมายข้อมูลและลงข้อสรุป

1. ในแต่ละข้อย่อย เช่น ข้อ 5.1 – 5.4 กำหนดน้ำหนักคะแนนข้อละ 1 คะแนน ถ้าตอบถูกต้องตามแนวทางในการตอบ หรือ มีส่วนใกล้เคียง หรือ มีแนวทางที่จะนำไปสู่คำตอบ โดยที่ผู้เข้าสอบจะต้องตอบข้อย่อยถูกทั้งสองข้อ เช่น ตอบข้อ 5.1.1 และ 5.1.2 ถูกทั้งสองข้อ จึงจะได้ 1 คะแนน ในข้อ 5.1 ถ้าไม่เป็นไปตามเงื่อนไขที่กำหนด ให้ 0 คะแนน

2. นำผลการตรวจให้คะแนนมาจัดระดับคุณภาพ โดยใช้กฎเกณฑ์การให้คะแนนทักษะการตีความหมายข้อมูลและลงข้อสรุปประกอบการพิจารณาพฤติกรรมที่บ่งชี้ความสามารถของผู้เข้าสอบ กำหนดคุณภาพเป็น 4 ระดับ บันทึกผลการให้คะแนนในภาพรวมของทักษะ

แนวทางในการตอบ

5.1.1 เปลี่ยนแปลง กล่าวคือ จำนวนปลาที่จับได้ในแต่ละปีลดน้อยลง ยกเว้น ปี พ.ศ. 2517 และ พ.ศ. 2519 ที่เพิ่มขึ้นเล็กน้อยจากปีก่อนหน้านั้น แต่กลับลดลงในปีถัดไป

5.1.2 ปริมาณปลาในทะเลบริเวณที่เคยจับได้ลดลงทุก ๆ ปี

5.2.1 ที่อุณหภูมิ 40 °C และ 50 °C สาร A ละลายในของเหลว B ได้แตกต่างกัน 20 กรัม

5.2.2 จากข้อมูลในตารางเมื่ออุณหภูมิสูงขึ้น สรุปการละลายของสาร A ได้ว่า ยิ่งอุณหภูมิสูงขึ้น สาร A ยิ่งละลายในของเหลว B ได้มากขึ้น

5.3.1 ความดันบรรยากาศที่ระดับความสูง 6 กิโลเมตร จากระดับน้ำทะเลมีค่าเท่ากับ 360 มิลลิเมตรปรอท

5.3.4 จากตารางข้อมูลข้างต้น สรุปความสัมพันธ์ระหว่างระดับความสูง ความดันบรรยากาศและความหนาแน่นของอากาศได้ว่า ยิ่งความสูงจากระดับน้ำทะเลเพิ่มมากขึ้น ความดันบรรยากาศและความหนาแน่นของอากาศมีค่าลดลง

5.4.1 พลังงานที่ใช้ในปี พ.ศ. 2505 กับปี พ.ศ. 2515 ต่างกันประมาณ 59,000 ล้านกิโลแคลอรี

5.4.2 จากข้อมูลในแผนภูมิข้างต้น สรุปความสัมพันธ์ระหว่างปริมาณพลังงานที่ใช้ในประเทศไทยกับระยะเวลาในทุก ๆ 2 ปีได้ว่า ปริมาณพลังงานที่ใช้ในประเทศไทยจากปี พ.ศ. 2501 ถึง พ.ศ. 2515 มีปริมาณเพิ่มมากขึ้นเรื่อย ๆ

กฎเกณฑ์การให้คะแนนทักษะการตีความหมายข้อมูลและลงข้อสรุป

ระดับคุณภาพ	คำอธิบาย
4 ดีมาก	หมายถึง แปลความหมายของข้อมูลที่จัดไว้ในรูปแบบต่าง ๆ ที่ใช้ในการสื่อความหมาย ข้อมูลในเชิงสถิติ พร้อมทั้งสรุปความสัมพันธ์ของข้อมูลได้ถูกต้องทั้งหมด
3 ดี	หมายถึง แปลความหมายของข้อมูลที่จัดไว้ในรูปแบบต่าง ๆ ที่ใช้ในการสื่อความหมาย ข้อมูลในเชิงสถิติ พร้อมทั้งสรุปความสัมพันธ์ของข้อมูลได้ถูกต้อง 3 ข้อใน 4 ข้อ
2 ใช้ได้	หมายถึง แปลความหมายของข้อมูลที่จัดไว้ในรูปแบบต่าง ๆ ที่ใช้ในการสื่อความหมาย ข้อมูลในเชิงสถิติ พร้อมทั้งสรุปความสัมพันธ์ของข้อมูลได้ถูกต้อง 2 ข้อใน 4 ข้อ
1 ใช้ไม่ได้	หมายถึง แปลความหมายของข้อมูลที่จัดไว้ในรูปแบบต่าง ๆ ที่ใช้ในการสื่อความหมาย ข้อมูลในเชิงสถิติ พร้อมทั้งสรุปความสัมพันธ์ของข้อมูลได้ถูกต้อง 1 ข้อใน 4 ข้อ
0	ไม่ตอบหรือตอบไม่ถูกเลย

ตาราง 21 ผลการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
ตามแบบแผน $i \times r \times p$

เลขที่	ผู้ตรวจคนที่ 1				
	S 1	S 2	S 3	S 4	S 5
1	3	2	3	3	3
2	3	2	4	3	3
3	2	2	2	2	2
4	3	2	3	3	3
5	2	1	2	2	2
6	3	2	3	3	3
7	3	2	3	2	3
8	3	2	2	2	3
9	3	2	3	2	3
10	2	2	2	2	2
11	2	2	2	2	2
12	2	2	2	2	2
13	3	3	2	3	3
14	2	2	2	2	2
15	3	2	2	2	3
16	3	2	2	2	3
17	3	2	1	2	3
18	2	2	2	2	2
19	3	3	2	3	3
20	4	3	3	3	4
21	2	2	3	2	2
22	2	3	2	2	2
23	2	2	2	2	2
24	2	2	2	2	2
25	2	2	2	2	2

เลขที่	ผู้ตรวจคนที่ 1				
	S 1	S 2	S 3	S 4	S 5
26	2	2	2	2	2
27	1	1	2	2	1
28	2	2	1	2	2
29	2	2	1	2	2
30	3	1	2	1	3
31	4	2	2	3	4
32	3	2	1	2	3
33	3	3	2	3	3
34	3	2	2	2	3
35	2	2	2	2	2
36	2	2	1	2	2
37	2	2	2	2	2
38	3	3	3	3	3
39	2	2	1	2	2
40	3	2	3	3	3
41	2	1	1	1	2
42	2	2	2	2	2
43	3	2	2	3	3
44	2	2	1	2	2
45	1	2	1	2	1
46	3	2	2	2	3
47	2	2	2	2	2
48	2	2	2	2	2
49	2	2	2	2	2
50	2	1	1	2	2

ตาราง 21 (ต่อ)

เลขที่	ผู้ตรวจคนที่ 2				
	S 1	S 2	S 3	S 4	S 5
1	2	3	3	3	3
2	2	4	3	3	3
3	2	2	2	2	2
4	2	3	3	3	3
5	1	2	2	2	2
6	2	3	3	3	3
7	2	3	2	2	3
8	2	2	2	2	3
9	2	3	2	2	3
10	2	2	2	2	2
11	2	2	2	2	2
12	2	2	2	2	2
13	3	2	3	3	3
14	2	2	2	2	2
15	2	2	2	2	3
16	2	2	2	2	3
17	2	1	2	2	3
18	2	2	2	2	2
19	3	2	3	3	3
20	3	3	3	3	4
21	2	3	2	2	2
22	3	2	2	2	2
23	2	2	2	2	2
24	2	2	2	2	2
25	2	2	2	2	2

เลขที่	ผู้ตรวจคนที่ 2				
	S 1	S 2	S 3	S 4	S 5
26	2	2	2	2	2
27	1	2	2	2	1
28	2	1	2	2	2
29	2	1	2	2	2
30	1	2	1	1	3
31	2	2	3	3	4
32	2	1	2	2	3
33	3	2	3	3	3
34	2	2	2	2	3
35	2	2	2	2	2
36	2	1	2	2	2
37	2	2	2	2	2
38	3	3	3	3	3
39	2	1	2	2	2
40	2	3	3	3	3
41	1	1	1	1	2
42	2	2	2	2	2
43	2	2	3	3	3
44	2	1	2	2	2
45	2	1	2	2	1
46	2	2	2	2	3
47	2	2	2	2	2
48	2	2	2	2	2
49	2	2	2	2	2
50	1	1	2	2	2

ตาราง 21 (ต่อ)

เลขที่	ผู้ตรวจคนที่ 3				
	S 1	S 2	S 3	S 4	S 5
1	3	3	3	3	2
2	4	3	3	3	2
3	2	2	2	2	2
4	3	3	3	3	2
5	2	2	2	2	1
6	3	3	3	3	2
7	3	2	2	3	2
8	2	2	2	3	2
9	3	2	2	3	2
10	2	2	2	2	2
11	2	2	2	2	2
12	2	2	2	2	2
13	2	3	3	3	3
14	2	2	2	2	2
15	2	2	2	3	2
16	2	2	2	3	2
17	1	2	2	3	2
18	2	2	2	2	2
19	2	3	3	3	3
20	3	3	3	4	3
21	3	2	2	2	2
22	2	2	2	2	3
23	2	2	2	2	2
24	2	2	2	2	2
25	2	2	2	2	2

เลขที่	ผู้ตรวจคนที่ 3				
	S 1	S 2	S 3	S 4	S 5
26	2	2	2	2	2
27	2	2	2	1	1
28	1	2	2	2	2
29	1	2	2	2	2
30	2	1	2	3	1
31	2	3	3	4	2
32	1	2	2	3	2
33	2	3	3	3	3
34	2	2	2	3	2
35	2	2	2	2	2
36	1	2	2	2	2
37	2	2	2	2	2
38	3	3	3	3	3
39	1	2	2	2	2
40	3	3	3	3	2
41	1	1	2	2	1
42	2	2	2	2	2
43	2	3	3	3	2
44	1	2	2	2	2
45	1	2	2	2	2
46	2	2	2	3	2
47	2	2	2	2	2
48	2	2	2	2	2
49	2	2	2	2	2
50	1	2	2	2	1

ตาราง 21 (ต่อ)

เลขที่	ผู้ตรวจคนที่ 4				
	S 1	S 2	S 3	S 4	S 5
1	3	3	3	2	3
2	4	3	3	4	3
3	2	2	2	2	2
4	3	3	3	2	3
5	2	2	2	1	2
6	3	3	3	2	3
7	3	2	3	2	2
8	2	2	3	2	2
9	3	2	3	2	2
10	2	2	2	2	2
11	2	2	2	2	2
12	2	2	2	2	2
13	2	3	3	3	3
14	2	2	2	2	2
15	2	2	3	2	2
16	2	2	3	2	2
17	1	2	3	2	2
18	2	2	2	2	2
19	2	3	3	3	3
20	3	3	4	3	3
21	3	2	2	2	2
22	2	2	2	3	2
23	2	2	2	2	2
24	2	2	2	2	2
25	2	2	2	2	2

เลขที่	ผู้ตรวจคนที่ 4				
	S 1	S 2	S 3	S 4	S 5
26	2	2	2	2	2
27	2	2	2	1	2
28	1	2	2	2	2
29	1	2	2	2	2
30	2	1	3	1	1
31	2	3	4	2	3
32	1	2	3	2	2
33	2	3	3	3	3
34	2	2	3	2	2
35	2	2	2	2	2
36	1	2	2	2	2
37	2	2	2	2	2
38	3	3	3	3	3
39	1	2	2	2	2
40	3	3	3	2	3
41	1	1	1	1	1
42	2	2	2	2	2
43	2	3	3	2	3
44	1	2	2	2	2
45	1	2	1	2	2
46	2	2	3	2	2
47	2	2	2	2	2
48	2	2	2	2	2
49	2	2	2	2	2
50	1	2	2	1	2

ตาราง 21 (ต่อ)

เลขที่	ผู้ตรวจคนที่ 5				
	S 1	S 2	S 3	S 4	S 5
1	3	3	3	3	2
2	3	3	3	3	2
3	2	2	2	2	2
4	3	3	3	3	2
5	2	2	2	2	1
6	3	3	3	3	2
7	2	2	3	3	2
8	2	2	3	3	2
9	2	2	3	3	2
10	2	2	2	2	2
11	2	2	2	2	2
12	2	2	2	2	2
13	3	3	3	3	3
14	2	2	2	2	2
15	2	2	3	3	2
16	2	2	3	3	2
17	2	2	3	3	2
18	2	2	2	2	2
19	3	3	3	3	3
20	3	3	4	4	3
21	2	2	2	2	2
22	2	2	2	2	3
23	2	2	2	2	2
24	2	2	2	2	2
25	2	2	2	2	2
26	2	2	2	2	2
27	2	2	1	1	1
28	2	2	2	2	2
29	2	2	2	2	2
30	1	1	3	3	2
31	3	3	4	4	2
32	2	2	3	3	2
33	3	3	3	3	3
34	2	2	3	3	2
35	2	2	2	2	2
36	2	2	2	2	2
37	2	2	2	2	2
38	3	3	3	3	3
39	2	2	2	2	2
40	3	3	3	3	2
41	1	1	2	2	1
42	2	2	2	2	2
43	3	3	3	3	2
44	2	2	2	2	2
45	2	2	2	2	2
46	2	2	3	3	2
47	2	2	2	2	2
48	2	2	2	2	2
49	2	2	2	2	2
50	2	2	2	2	1

ตาราง 22 ผลการตรวจให้คะแนนแบบทดสอบวัดทักษะกระบวนการทางวิทยาศาสตร์ชั้นบูรณาการ
ตามแบบแผน $i \times (p:r)$

เลขที่	ผู้ตรวจคนที่ 1				
	S 1	S 2	S 3	S 4	S 5
1	3	3	1	3	3
2	2	2	2	3	2
3	3	3	2	2	2
4	2	2	2	3	2
5	3	3	2	3	3
6	2	2	2	3	2
7	3	3	2	2	2
8	2	2	2	3	2
9	2	2	2	3	2
10	3	3	2	3	3

เลขที่	ผู้ตรวจคนที่ 2				
	S 1	S 2	S 3	S 4	S 5
11	3	3	2	2	2
12	2	2	2	3	2
13	2	2	2	3	2
14	3	3	3	3	3
15	2	2	2	3	2
16	3	3	2	2	2
17	2	2	2	3	2
18	3	3	2	3	3
19	2	2	2	3	2
20	2	2	2	3	2

เลขที่	ผู้ตรวจคนที่ 3				
	S 1	S 2	S 3	S 4	S 5
21	2	2	2	3	2
22	2	2	2	3	2
23	3	3	2	3	3
24	3	3	2	3	3
25	2	2	2	1	2
26	3	1	2	1	3
27	2	2	2	2	2
28	2	2	2	2	2
29	2	2	2	2	2
30	3	2	2	2	3

เลขที่	ผู้ตรวจคนที่ 4				
	S 1	S 2	S 3	S 4	S 5
31	2	2	2	2	2
32	2	2	2	2	2
33	2	3	2	2	2
34	2	2	2	2	2
35	3	2	2	2	3
36	2	2	4	2	2
37	2	2	2	2	2
38	2	2	2	2	2
39	3	2	2	2	3
40	2	2	2	2	2

ตาราง 22 (ต่อ)

เลขที่	ผู้ตรวจคนที่ 5				
	S 1	S 2	S 3	S 4	S 5
41	2	2	2	2	2
42	2	2	2	2	2
43	3	2	2	2	3
44	2	3	2	2	2
45	2	3	2	1	2
46	2	2	2	2	2
47	2	2	2	2	2
48	3	2	2	2	3
49	3	2	2	2	3
50	2	2	1	4	2

หมายเหตุ	S 1	หมายถึง	ทักษะการตั้งสมมติฐาน
	S 2	หมายถึง	ทักษะการกำหนดนิยามเชิงปฏิบัติการ
	S 3	หมายถึง	ทักษะการกำหนดและควบคุมตัวแปร
	S 4	หมายถึง	ทักษะการทดลอง
	S 5	หมายถึง	ทักษะการตีความหมายข้อมูลและลงข้อสรุป

ประวัติย่อผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ ชื่อสกุล	นายชาญวิทย์ จรัสสุทธิอิศร
วันเดือนปีเกิด	8 มีนาคม 2502
สถานที่เกิด	อำเภอราษีไศล จังหวัดศรีสะเกษ
สถานที่อยู่ปัจจุบัน	381/52 ถนนท้าวสุระ (หมู่บ้านพลลำนวิลา) ต.ในเมือง อ.เมืองฯ จ. นครราชสีมา 30000
ตำแหน่งหน้าที่การงานในปัจจุบัน	ศึกษานิเทศก์ 7 หน่วยศึกษานิเทศก์ กรมสามัญศึกษา เขตการศึกษา 11 จ.นครราชสีมา
สถานที่ทำงานปัจจุบัน	ฝ่ายนิเทศงานแผนงานและวิจัย หน่วยศึกษานิเทศก์ กรมสามัญศึกษา เขตการศึกษา 11 ถนนสีปรี ต.ในเมือง อ.เมืองฯ จ.นครราชสีมา 30000
ประวัติการศึกษา	
พ.ศ. 2520	ม.ศ. 5 (แผนกวิทย์-คณิต) โรงเรียนศรีสะเกษวิทยาลัย
พ.ศ. 2522	ป.กศ. ชั้นสูง (วิทยาศาสตร์) วิทยาลัยครูอุบลราชธานี
พ.ศ. 2528	ศษ.บ. (วิทยาศาสตร์) มหาวิทยาลัยสุโขทัยธรรมาธิราช
พ.ศ. 2545	กศ.ม. (การวัดผลการศึกษา) มหาวิทยาลัยศรีนครินทรวิโรฒ