

371.2.013

๒๕๔๑๑

ร.๓

ความเชื่อมั่นของการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนต่างกัน

บทคัดย่อ

ของ

ชัยยุทธ คชรินทร์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา

พฤษภาคม 2546

๒๕๔๑๑ ร.๓

ชัยยุทธ คชรินทร์. (2546), ความเชื่อมั่นของการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนต่างกัน.

ปริญญาานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัย

ศรีนครินทรวิโรฒ. คณะกรรมการควบคุม : รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนตพงษ์ ,

อาจารย์ชวลิต รวยอาจิณ

การวิจัยครั้งนี้ มีจุดมุ่งหมายสำคัญเพื่อศึกษาและเปรียบเทียบค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) คำนวณด้วยสูตร P_0 ของแฮมเบลตันและโนวิก, สูตร K ของสก็อต และสูตร RAI ของเบอร์รี่ – สติอก และคณะ ผู้ประเมินเป็นครูผู้สอนวิชาศิลปะศึกษาที่สังกัดสำนักงานเขตคลองสามวา กรุงเทพมหานคร

ผลการวิจัยพบว่า

1. ค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) สูตร P_0 ให้ค่าความเชื่อมั่นสูงเป็นอันดับแรก รองลงมาคือ สูตร RAI และสูตร K ตามลำดับ และเมื่อใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม (Holistic Scoring Rubric) สูตร RAI ให้ค่าความเชื่อมั่นสูงเป็นอันดับแรก รองลงมาคือ สูตร K และสูตร P_0 ตามลำดับ

2. สูตร P_0 , สูตร K และสูตร RAI ให้ค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งที่คำนวณจากคะแนนการประเมินที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

3. คะแนนการประเมินที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) เมื่อคำนวณด้วยสูตร P_0 ให้ค่าความเชื่อมั่นแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ .01 แต่คำนวณด้วยสูตรอื่น ให้ค่าความเชื่อมั่นไม่แตกต่าง

RELIABILITY OF ARTWORK RATING USING DIFFERENT SCORING RUBRIC

AN ABSTRACT
BY
CHAIYUT COTECHARIN

Presented in partial fulfillment of the requirements
for the Master of Education degree in Educational Measurement
at Srinakharinwirot University
May 2003

Chaityut Cotecharin. (2003) The Reliability Of Artwork Rating Using Different Scoring Rubric.
Master Thesis, M.Ed. (Educational Measurement) Bangkok : Graduate School,
Srinakharinwirot University. Advisor Committee : Assoc. Prof. Dr. Boonchird
Pinyoanuntapong and Mr.Chavalit Ruayajin.

This study aimed to examine and compare the reliability of artwork rating using different scoring rubric by Analytic Scoring Rubric and Holistic Scoring Rubric calculated by P_0 formula of Hambleton and Novick, K formula of Scott, and RAI formula of Burry – Stock and others. Art teachers under office of Klong Samwa district, Bangkok. Metropolitan were the evaluators.

The results revealed that :

1. The reliability of artwork rating using Analytic Scoring Rubric with P_0 formula, RAI and K formula were at a high level respectively, when using Holistic Scoring Rubric, it was found that RAI formula got highest reliability and then K and P_0 formula respectively.
2. There was no significant difference among the reliability of artwork rating using Analytic Scoring Rubric and Holistic Scoring Rubric with P_0 formula, RAI and K formula.
3. There was a significant difference at .01 level among score rating using Analytic Scoring Rubric and Holistic Scoring Rubric calculated by P_0 formula, whereas there was no significant difference among those with other formulas.

ความเชื่อมั่นของการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนต่างกัน

ปริญญานิพนธ์
ของ
ชัยยุทธ คชรินทร์

เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา
พฤษภาคม 2546
ลิขสิทธิ์เป็นของมหาวิทยาลัยศรีนครินทรวิโรฒ

ปริญญานิพนธ์

เรื่อง

ความเชื่อมั่นของการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนต่างกัน

ของ

นายชัยยุทธ คชรินทร์

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้เป็นส่วนหนึ่งของการศึกษาหลักสูตร

ปริญญาการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษา

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....คณบดีบัณฑิตวิทยาลัย

(รองศาสตราจารย์ ดร. นภาพร หะวานนท์)

วันที่ เดือน พ.ศ. 2546

คณะกรรมการสอบปริญญานิพนธ์

.....ประธาน

(รองศาสตราจารย์ ดร. บุญเชิด ภิญโญนนท์พงษ์)

.....กรรมการ

(อาจารย์ ขวลิต รวยอาจิน)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(รองศาสตราจารย์ ดร. ส. วาสนา ประवालพฤกษ์)

.....กรรมการที่แต่งตั้งเพิ่มเติม

(อาจารย์ ดร. สุพร เข้มเฮง)

ประกาศคุณูปการ

ปริญญานิพนธ์ฉบับนี้สำเร็จสมบูรณ์ได้ เพราะได้รับความกรุณาจาก รองศาสตราจารย์ ดร. บุญเชิด ภิญโญอันตพงษ์ ประธานคณะกรรมการควบคุมและ อาจารย์ชวลิต รวยอาจัน กรรมการให้คำปรึกษา แนะนำ ชี้นำ และตรวจสอบ งานวิจัยนี้จนสามารถบรรลุจุดมุ่งหมาย ข้าพเจ้าขอกราบขอบพระคุณท่านเป็นอย่างสูง

ขอกราบขอบพระคุณ รองศาสตราจารย์ ดร. ส.วาสนา ประवालพฤษ์ และ ดร. สุวพร เข้มเฮง กรรมการที่ แต่งตั้งเพิ่มเติม ที่ได้ให้คำแนะนำ ชี้นำ และปรับปรุงรูปเล่ม เนื้อหาสาระที่ถูกต้องชัดเจนยิ่งขึ้น เพื่อให้งานวิจัยนี้มีคุณภาพที่ดี ข้าพเจ้าขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ คณะจารย์ ของภาควิชาวัดผลและวิจัยการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ ทุกท่านที่ได้ช่วยดูแล ให้คำปรึกษาแนะนำแนวทางในการทำวิจัยตลอดมา ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบขอบพระคุณ อาจารย์ชาญวิทย์ จรัสสุทธิศิคร อาจารย์นิศารัตน์ คงสวัสดิ์ อาจารย์ปวีดา ชื่นชัย อาจารย์ไพฑูรย์ ภูมิซ้อ อาจารย์ธงชัย ดวงมุลตรี ที่ให้ความอนุเคราะห์ช่วยเหลือในการตรวจสอบ เครื่องมือการวิจัย

ขอกราบขอบพระคุณ อาจารย์นุชสราร กิตติสารโร อาจารย์ประสาศ ยินดี อาจารย์ธงชัย ดวงมุลตรี ที่ให้ความอนุเคราะห์ช่วยเหลือในการพัฒนาเครื่องมือการวิจัย

ขอกราบขอบพระคุณนายขวัญชัย บุญคุ้ม ผู้บริหารโรงเรียนสุเหร่าคลองหนึ่ง อาจารย์โรงเรียนสุเหร่าคลองหนึ่งทุกท่าน ที่ให้กำลังใจ ด้วยกันตลอดมา

ขอกราบขอบพระคุณ อาจารย์สุเทพ อินทกุล อาจารย์วิทยา ดันตัยย์ อาจารย์มานพ มินยา อาจารย์ภัทรพงศ์ สุดตะ และอาจารย์ผดุงศักดิ์ ศัพทเสน ที่ให้ความช่วยเหลือเป็นผู้ประเมินผลงานศิลปะ เพื่อนำผลมาวิเคราะห์ข้อมูล การวิจัย

ขอขอบคุณสมาชิกในครอบครัว และเพื่อนๆ ผู้ร่วมงานทุกท่านที่ให้กำลังใจ และช่วยเหลืออย่างดี ตลอดระยะเวลาที่ศึกษาและทำการวิจัย และขอบคุณเพื่อนๆ นิสิตระดับปริญญาโท สาขาวิชาการวัดผลการศึกษา ภาคพิเศษ รุ่นที่ 11 ทุกคนที่ให้กำลังใจ ร่วมมือค้นคว้าข้อมูลด้วยกันตลอดมา

คุณค่าและประโยชน์ ๆ ที่ได้รับจากปริญญานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องบูชาพระคุณ บิดา มารดา ครูอาจารย์ และผู้มีพระคุณทุกท่าน

ชัยยุทธ กชรินทร์

สารบัญ

บทที่	หน้า
1 บทนำ.....	1
ภูมิหลัง.....	1
ความมุ่งหมายของการวิจัย.....	2
ความสำคัญของการวิจัย.....	2
ขอบเขตของการวิจัย.....	3
ผู้ประเมินที่ใช้ในการวิจัย.....	3
ตัวแปรที่ศึกษา.....	3
นิยามศัพท์เฉพาะ.....	3
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	5
เอกสารที่เกี่ยวข้องกับการวิจัย.....	5
ความเชื่อมั่น.....	5
ความหมายของความเชื่อมั่น.....	5
ทฤษฎีความเชื่อมั่น.....	6
ความเชื่อมั่นของผู้ประเมิน.....	8
สูตรที่ใช้คำนวณหาค่าความเชื่อมั่นของการประเมินผลงาน.....	11
การวัดภาคปฏิบัติ.....	13
ความหมายของการวัดภาคปฏิบัติ.....	13
พฤติกรรมการเรียนรู้ทางภาคปฏิบัติ.....	15
เกณฑ์ที่ใช้ในการพิจารณาตัดสินเกี่ยวกับการวัดภาคปฏิบัติ.....	17
ชนิดของแบบทดสอบวัดภาคปฏิบัติ.....	18
เครื่องมือวัดภาคปฏิบัติ.....	20
การตรวจให้คะแนนการทดสอบภาคปฏิบัติ.....	24
เกณฑ์การให้คะแนนผลงาน.....	25
ประเภทของกฎเกณฑ์การให้คะแนน.....	26
การสร้างกฎเกณฑ์การให้คะแนน.....	27
แนวทางในการตรวจให้คะแนน.....	32
ลักษณะของศิลปะ.....	36
งานวิจัยที่เกี่ยวข้อง.....	42
งานวิจัยในต่างประเทศ.....	42
งานวิจัยในประเทศ.....	44
สมมติฐานในการวิจัย.....	47

สารบัญ(ต่อ)

บทที่	หน้า
3	วิธีดำเนินการวิจัย..... 48
	การกำหนดประชากรและเลือกกลุ่มตัวอย่าง..... 48
	ผู้ประเมินที่ใช้ในการวิจัย..... 48
	เครื่องมือที่ใช้ในการวิจัย..... 48
	การสร้างเครื่องมือที่ใช้ในการวิจัย..... 49
	วิธีดำเนินการเก็บรวบรวมข้อมูล..... 55
	การจัดกระทำและการวิเคราะห์ข้อมูล..... 56
4	ผลการวิเคราะห์ข้อมูล..... 60
	สัญลักษณ์และอักษรย่อในการวิเคราะห์ข้อมูล..... 60
	ผลการวิเคราะห์ข้อมูล..... 60
5	สรุปผล อภิปราย และข้อเสนอแนะ..... 64
	ความมุ่งหมายของการวิจัย..... 64
	สมมติฐานของการวิจัย..... 64
	ผู้ประเมินที่ใช้ในการวิจัย..... 65
	เครื่องมือที่ใช้ในการวิจัย..... 65
	วิธีดำเนินการเก็บรวบรวมข้อมูล..... 65
	สรุปผลการวิเคราะห์ข้อมูล..... 67
	อภิปรายผล..... 67
	ข้อเสนอแนะ..... 68
บรรณานุกรม.....	69
ภาคผนวก.....	74
ประวัติย่อผู้วิจัย.....	96

บัญชีตาราง

ตาราง		หน้า
1	ผลรวมของการประเมินให้คะแนนแบบ 2 ขั้ว จากผู้ประเมิน 2 คน.....	9
2	ผลรวมของการให้คะแนนแบบหลายขั้ว โดยผู้ประเมิน 2 คน.....	10
3	เกณฑ์การให้คะแนนแบบรวมองค์ประกอบ.....	28
4	เกณฑ์การให้คะแนนแบบแยกองค์ประกอบ.....	29
5	คำอธิบายคุณภาพ.....	30
6	การกำหนดคะแนนการตัดสินระดับคุณภาพ.....	31
7	จำนวนผู้ประเมินจำแนกตามโรงเรียน.....	48
8	ค่าสถิติพื้นฐานของคะแนนผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนแบบเกณฑ์แยกและ เกณฑ์รวมของผู้ประเมิน.....	61
9	ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่คำนวณด้วยสูตรต่างกัน.....	62
10	การเปรียบเทียบค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่มีสูตรต่างกัน.....	62
11	การเปรียบเทียบค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่มีรูปแบบประเมินต่างกัน.....	63
12	ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือของการให้คะแนนการประเมิน แบบเกณฑ์แยก.....	78
13	ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือของการให้คะแนนการประเมิน แบบเกณฑ์รวม.....	78
14	ค่าอำนาจจำแนกของมาตรประเมินผลงานศิลปะ.....	79
15	เปรียบเทียบค่าสถิติพื้นฐาน การประเมินด้วยการให้คะแนนการประเมินแบบเกณฑ์แยกและ แบบเกณฑ์รวม.....	88
16	คะแนนการประเมินผลงานศิลปะ เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยกและ ค่าสถิติพื้นฐาน.....	89
17	คะแนนการประเมินผลงานศิลปะ เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวมและ ค่าสถิติพื้นฐาน.....	91
18	ค่าความเชื่อมั่นสูตร P_0 ของแสมเบิลตัน และโนวิก เมื่อใช้การให้คะแนนการประเมิน แบบเกณฑ์แยก.....	92
19	ค่าความเชื่อมั่นสูตร P_0 ของแสมเบิลตัน และโนวิก เมื่อใช้การให้คะแนนการประเมิน แบบเกณฑ์รวม.....	92
20	ค่าความเชื่อมั่นสูตร K ของสก็อต เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก.....	93
21	ค่าความเชื่อมั่นสูตร K ของสก็อต เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวม.....	93
22	ค่าความเชื่อมั่นสูตร RAI ของเบอร์รี – สต็อก และคณะ เมื่อใช้การให้คะแนนการประเมิน แบบเกณฑ์แยก.....	94
23	ค่าความเชื่อมั่นสูตร RAI ของเบอร์รี – สต็อก และคณะ เมื่อใช้การให้คะแนนการประเมิน แบบเกณฑ์รวม.....	94

บัญชีภาพประกอบ

ภาพประกอบ	หน้า
1 แสดงลำดับขั้นการสร้างแบบประเมินผลงานศิลปะ.....	49
2 แสดงลำดับขั้นการสร้างกฎเกณฑ์การให้คะแนนการประเมินผลงานศิลปะ.....	51

บทที่ 1

บทนำ

ภูมิหลัง

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 หมวด 4 การจัดการศึกษา มาตรา 26 ระบุไว้ว่าให้สถานศึกษาจัดการประเมินผู้เรียน โดยพิจารณาจากพัฒนาการของผู้เรียน ความประพฤติ การสังเกตพฤติกรรมการเรียน การร่วมกิจกรรม และการทดสอบควบคู่ไปในกระบวนการเรียนการสอน ให้มีวิธีการประเมินที่หลากหลาย (พ.ร.บ.การศึกษาแห่งชาติ พ.ศ. 2542. 2542 :12 – 14) ให้มีความเหมาะสมกับวัยและลักษณะความสามารถของเด็ก ดังนั้นในการประเมินนักเรียนจึงจำเป็นต้องใช้เครื่องมือชนิดต่างๆ ทั้งการวัดแบบเป็นทางการและไม่เป็นทางการ

ศิลปศึกษาเป็นวิชาหนึ่งที่มีความสำคัญไม่น้อยเช่นเดียวกับวิชาอื่นๆ การเรียนศิลปศึกษาเป็นการฝึกความคิดให้สัมพันธ์กับมือและความคิดสร้างสรรค์ รู้จักแก้ปัญหาในการทำงาน หัดให้มีความเป็นระเบียบ รู้จักรักสวย รักงาม มีสุนทรีย์ในการรับรู้ เมื่อพบเห็นสิ่งต่าง ๆ มีความซาบซึ้ง (Appreciate) ในผลงานศิลปะ สามารถสร้างผลงานต่าง ๆ ได้ด้วยตนเอง ทำให้มีความเชื่อมั่นและความภาคภูมิใจ เข้าใจศิลปวัฒนธรรมของโลก ทำให้เกิดความรัก และหวงแหนศิลปกรรมของชาติและก่อให้เกิดการอนุรักษ์นอกจากนี้ ศิลปะยังช่วยผ่อนคลายความเครียด ช่วยในการประหยัด และเป็นพื้นฐานอาชีพได้เป็นอย่างดี (เกษร ชิตะจาวี. 2543)

การประเมินผลแบบเดิมของงานศิลปะ ครูจะนำผลงานของนักเรียนทั้งชั้นมาจัดวางเรียงลำดับความสามารถจากพวกเก่งที่สุดไปจนถึงอ่อนที่สุด โดยนำผลงานทั้งหมดเปรียบเทียบกัน พิจารณาตามระดับมาตรฐานและความพอใจของครู การตัดสินชี้ขาดให้คะแนนคำนึงถึงสัดส่วนความถูกต้อง หลักการจัดภาพ ทฤษฎีสี ฯลฯ และผลงานศิลปะที่สร้างขึ้นในแต่ละครั้งจะมีลักษณะที่เหมือนๆ กันทั้งชั้น วิธีการสอนก็ปิดกั้นการแสดงออกทางความคิดริเริ่มสร้างสรรค์และทำลายทัศนคติทางศิลปะของนักเรียนโดยสิ้นเชิง โดยเฉพาะนักเรียนในระดับประถมศึกษา

การให้คะแนนนิยมเขียนให้เป็นตัวเลข เช่น คะแนนเต็ม 10 ดีมาก, ดีให้ 7, 8, 9 ปานกลางให้ 4, 5, 6 อ่อน, อ่อนมาก ยังต้องปรับปรุงแก้ไขให้ 1, 2, 3 หรืออีกวิธีหนึ่ง ก็คือ การเขียนให้ความเห็นว่า ดี, พอใช้, พอใช้ได้ เป็นต้น (เลิศ อานันท์นะ. 2523) และผู้ประเมินที่ไม่มีความรู้ย่อมไม่ทราบว่าจะวัดองค์ประกอบใดบ้าง รายการพฤติกรรมที่ควรจะนำมาใช้วัดมีอะไรบ้าง การให้น้ำหนักความสำคัญแก่องค์ประกอบต่าง ๆ จะต้องปฏิบัติอย่างไร ควรให้น้ำหนักการวัดที่กระบวนการ (Process) หรือผลงาน (Product) หรือทั้งสองส่วน การออกแบบเครื่องมือจึงไม่ได้มาตรฐาน การวัดให้คะแนนจึงไม่มีระบบ ไม่มีผู้ประเมินเป็นหลักในการตัดสินใจ สิ่งนี้ได้ส่งผลต่อไปถึงปัญหาในการกำหนดมาตรฐานการประเมิน เพราะผู้ประเมินแต่ละคนมีความแตกต่างกันในการใช้หลักการตัดสินใจ เครื่องมือวัดการปฏิบัตินั้นมักไม่ค่อยได้มีการตรวจสอบคุณภาพก่อนนำไปใช้หรือหลังจากนำไปใช้แล้ว (จินตนา ธนวิบูลย์ชัย. 2535)

ดังนั้นในการประเมินผลงานศิลปะ จึงต้องพัฒนาวิธีการประเมินให้คะแนนควบคู่กันไปด้วย และต้องตรวจสอบความเชื่อมั่นของการประเมินให้คะแนน สำหรับวิธีการคำนวณค่าความเชื่อมั่นของการประเมินมีหลายวิธีด้วยกัน แต่ละวิธีจะมีจุดเด่นจุดด้อยในตัวทั้งสิ้น การพิจารณาเลือกใช้วิธีใดขึ้นอยู่กับความง่าย และใช้ได้สะดวก จากการใช้สูตรการหาค่าความเชื่อมั่นจำนวน 3 สูตร คือ สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick. 1973) สูตร K ของสก๊อต (Scott's, 1971) และสูตร RAI ของเบอร์รี่ – สต็อกซ์ และคณะ (Judith A. Berry – Stock,

Cecelia, Brad S. Chissom) ซึ่งเป็นสูตรการหาค่าความเชื่อมั่นของการประเมินที่ไม่สลับซับซ้อนและน่าสนใจ ดังนั้นผู้วิจัยจึงสนใจนำสูตรนี้มาศึกษาค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่ใช้สูตรต่างกัน เพื่อจะได้เป็นแนวทางในการหาค่าความเชื่อมั่นของการประเมินผลงานศิลปะ โดยตั้งเป็นคำถามของการวิจัยว่า สูตรใดจะให้ค่าความเชื่อมั่นสูงกว่ากัน เมื่อมีวิธีการที่เหมือนกัน

ความมุ่งหมายของการวิจัย

การศึกษาวิจัยครั้งนี้ มีจุดมุ่งหมายสำคัญ เพื่อหาค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่มีกฎเกณฑ์การให้คะแนนต่างกัน จะมีค่าความเชื่อมั่นที่ต่างกันหรือไม่ เมื่อคำนวณจากสูตรที่ต่างกัน โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อหาค่าความเชื่อมั่นของการประเมินผลงานศิลปะ เมื่อคำนวณด้วยสูตรและกฎเกณฑ์การให้คะแนนการประเมินต่างกันดังนี้

1.1 ค่าความเชื่อมั่นที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

1.2 ค่าความเชื่อมั่นที่คำนวณด้วยสูตร K ของสก็อต (Scott's, 1971) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

1.3 ค่าความเชื่อมั่นที่คำนวณด้วยสูตร RAI ของเบอร์รี่ - สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

2. เพื่อเปรียบเทียบค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) สูตร K ของสก็อต (Scott's, 1971) และสูตร RAI ของเบอร์รี่ - สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

ความสำคัญของการวิจัย

ผลการศึกษาวิจัยนี้มีประโยชน์ต่อการศึกษาตามแนวพระราชบัญญัติการศึกษาแห่งชาติ พุทธศักราช 2542 ที่ต้องการให้มีการวัดและประเมินผลของนักเรียนในระหว่างเรียนอย่างแท้จริง ครูผู้สอนสามารถนำแบบประเมินผลงานในการวิจัยนี้ไปประยุกต์ใช้ในการประเมินผลงานของวิชาที่ตนเองสอนและมีแนวทางในการตรวจสอบคุณภาพการประเมินของผู้ประเมินด้วยวิธีง่าย ๆ โดยมีต้องมีความรู้ในเชิงสถิติขั้นสูง ดังนั้น จึงสะดวกต่อการใช้งานเป็นอย่างดี

ขอบเขตของการวิจัย

ผู้ประเมินที่ใช้ในการวิจัย

ผู้ประเมิน คือ ครูผู้สอนศิลปศึกษาที่จบปริญญาตรีด้านศิลปศึกษาและมีประสบการณ์การสอนด้านศิลปศึกษาอย่างน้อย 5 ปี สังกัดโรงเรียนในสำนักงานเขตคลองสามวา ซึ่งคัดเลือกมาจำนวน 5 คนจากผู้สมัครใจ มาทำการประเมินผลงานศิลปะของนักเรียน จำนวน 30 ภาพ

ตัวแปรที่ศึกษา

1. ตัวแปรอิสระ ในการศึกษาครั้งนี้ คือ

1.1 สูตรการหาค่าความเชื่อมั่นแบ่งเป็นสามสูตร ได้แก่

สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973)

สูตร K ของสก็อต (Scott's, 1971)

สูตร RAI ของเบอร์รี – สต็อกก์ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom)

1.2 กฎเกณฑ์การให้คะแนนการประเมินแบ่งเป็นสองแบบได้แก่

กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric)

กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม (Holistic Scoring Rubric)

2. ตัวแปรตาม ได้แก่ ค่าความเชื่อมั่นของการประเมินผลงานศิลปะ

นิยามศัพท์เฉพาะ

1. แบบประเมินผลงานศิลปะ หมายถึง เครื่องมือที่ใช้ในการประเมินภาพศิลปะตามเกณฑ์การวาดภาพระบายสีทั้งหมด 7 ประเด็น และมีกฎเกณฑ์การให้คะแนนการประเมินสองแบบ คือ แบบเกณฑ์แยกและแบบเกณฑ์รวม

2. กฎเกณฑ์การพิจารณาให้คะแนนความสามารถในการวาดภาพ ประกอบด้วย

2.1 การเน้นจุดเด่นของภาพ หมายถึง ภาพเด่นสะดุดตาและมีการใช้สีสะดุดตาชัดเจน

2.2 ความสมดุลของภาพ หมายถึง ภาพมีการจัดวางภาพได้เหมาะสมกับหน้ากระดาษและภาพมีระยะใกล้ - ไกลมีการผสมผสานกลมกลืนของรูปและพื้นหลัง

2.3 การระบายสี หมายถึง ภาพมีการระบายสีภาพอยู่ในกรอบสะอาดเรียบร้อย สีเรียบ กลมกลืน

2.4 การใช้สี หมายถึง ภาพให้แสงและเงาถูกต้อง และแยกความสัมพันธภาพ และพื้นหลังชัดเจน ให้สีเหมือนธรรมชาติ

2.5 ความคิดสร้างสรรค์ หมายถึง ภาพมีแนวคิดแปลกใหม่มีการสื่อความหมายของการใช้ความคิดของตนเองและไม่ซ้ำภาพของผู้อื่น

2.6 ความประณีต หมายถึง ภาพมีความละเอียดมีความชัดเจนของภาพและสามารถมองเห็นความงดงามของภาพ

2.7 ความสะอาด หมายถึง ภาพมีการระบายสีอยู่ในกรอบ ไม่เลอะเทอะและมีความเป็นระเบียบเรียบร้อย

3. กฎเกณฑ์การให้คะแนนการประเมิน หมายถึง ข้อกำหนดที่ใช้เป็นแนวทางในการให้คะแนนผลงานศิลปะ ซึ่งแบ่งเป็นสองแบบ คือ แบบเกณฑ์แยกและแบบเกณฑ์รวม

3.1 กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก หมายถึง การให้คะแนนการประเมินภาพ โดยให้คะแนนแยกทีละประเด็นเป็นระดับ 4, 3, 2, 1 ทั้งหมด 7 ประเด็น

3.2 กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม หมายถึง การให้คะแนนการประเมินในภาพรวมของ 7 ประเด็น เป็นคะแนนระดับ 4, 3, 2, 1 เพียงค่าเดียว

4. ผลงานศิลปะ หมายถึง ภาพวาดระบายสีของนักเรียนชั้นประถมศึกษาปีที่ 6 ของนักเรียนโรงเรียนสุเหร่าคลองหนึ่ง สังกัดสำนักงานเขตคลองสามวา กรุงเทพมหานคร จำนวน 30 ภาพ ที่คัดเลือกมาโดยครูผู้สอนวิชาศิลปะศึกษา

5. ความเชื่อมั่นของการประเมิน หมายถึง ผลการให้คะแนนการประเมินผลงานศิลปะของผู้ประเมิน ซึ่งคำนวณได้จาก สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) สูตร K ของสก็อต (Scott's, 1971) และสูตร RAI ของเบอร์รี่ – สต็อก และคณะ (Judith A. Berry – Stock, Cecelia, Brad S. Chissom)

6. ผู้ประเมิน หมายถึง ครูผู้สอนศิลปะศึกษาที่จบปริญญาตรีด้านศิลปะศึกษาและมีประสบการณ์การสอนด้านศิลปะศึกษาอย่างน้อย 5 ปี สังกัดโรงเรียนในสำนักงานเขตคลองสามวา ซึ่งคัดเลือกมาจำนวน 5 คน จากผู้สมัครใจ มาทำการประเมินผลงานศิลปะของนักเรียน จำนวน 30 ภาพ

7. ผู้เชี่ยวชาญทางวัดผล หมายถึง อาจารย์ผู้มีประสบการณ์ทางการวัดผล หรือผู้สำเร็จการศึกษาระดับปริญญาโทด้านการวัดผลการศึกษา ทำการตรวจสอบคุณภาพของแบบประเมินแบบเกณฑ์แยกและแบบเกณฑ์รวม จำนวน 2 ท่าน

8. ผู้เชี่ยวชาญทางศิลปะศึกษา หมายถึง อาจารย์ผู้มีประสบการณ์ทางการสอนศิลปะศึกษาอย่างน้อย 5 ปี หรือผู้สำเร็จระดับปริญญาตรี - ปริญญาโทด้านศิลปะศึกษา ทำการตรวจสอบคุณภาพของแบบประเมินแบบเกณฑ์แยกและแบบเกณฑ์รวม จำนวน 3 ท่าน

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

เอกสารและงานวิจัยที่เกี่ยวข้องกับการศึกษาค้นคว้า ผู้วิจัยได้ศึกษาเนื้อหาต่าง ๆ ดังนี้

1. ความเชื่อมั่น (Reliability)
 - 1.1 ความหมายของความเชื่อมั่น
 - 1.2 ทฤษฎีความเชื่อมั่น
 - 1.3 ความเชื่อมั่นของผู้ประเมิน
 - 1.4 สูตรที่ใช้คำนวณหาความเชื่อมั่นของผู้ประเมิน
2. การวัดภาคปฏิบัติ
 - 2.1 ความหมายของการวัดภาคปฏิบัติ
 - 2.2 พฤติกรรมการเรียนรู้ทางภาคปฏิบัติ
 - 2.3 เกณฑ์ที่ใช้ในการพิจารณาตัดสินเกี่ยวกับการวัดภาคปฏิบัติ
 - 2.4 ชนิดของแบบทดสอบวัดภาคปฏิบัติ
 - 2.5 เครื่องมือวัดภาคปฏิบัติ
 - 2.6 การตรวจให้คะแนนการทดสอบภาคปฏิบัติ
 - 2.7 เกณฑ์การให้คะแนนผลงาน
 - 2.8 ประเภทของกฎเกณฑ์การให้คะแนน
 - 2.9 การสร้างกฎเกณฑ์การให้คะแนน
 - 2.10 แนวทางในการตรวจให้คะแนน
 - 2.11 ลักษณะของศิลปะ
3. เอกสารและงานวิจัยที่เกี่ยวข้อง

1. ความเชื่อมั่น (Reliability)

1.1 ความหมายของความเชื่อมั่น

นักจิตวิทยาและนักการศึกษาได้ให้ความหมายของ ความเชื่อมั่นไว้ต่างกัน ดังนี้

นันทอลลี (Nunnally. 1964 : 59) กล่าวว่าความเชื่อมั่นเป็นสัดส่วนระหว่างความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่ได้จากการทดสอบ

อนาสตาซี (Anastasi. 1968 : 105) กล่าวว่าความเชื่อมั่นเป็นความคงที่ของคะแนนที่ได้จากการสอบบุคคลเดียวกันแต่ต่างเวลา และโอกาสกัน

ลอร์ด และโนวิก (Lord and Novick, 1968 : 46) กล่าวว่าความเชื่อมั่นเป็นความคงที่ของคะแนนที่ได้จากการตอบซ้ำ และคะแนนที่ได้จากการตอบแบบทดสอบทั้งสองครั้งเป็นอิสระไม่ขึ้นกับความคลาดเคลื่อนของการวัดใด ๆ

กรอนลันด์ (Gronlund, 1976 : 105) กล่าวว่า ความเชื่อมั่นเป็นค่าความคงที่ของคะแนนในการทดสอบหรือความคงที่จากการประเมินครั้งแรก และครั้งอื่น ๆ

ลินด์วอลล์และนิกโค (Lindvall and Nitko, 1967 : 126) กล่าวว่า ความเชื่อมั่นเป็นค่าสหสัมพันธ์ระหว่างคะแนนที่ได้จากการสอบสองครั้ง โดยใช้แบบทดสอบฉบับเดียวกัน สอบในเวลาต่างกัน

บุญเชิด ภิญญอนันตพงษ์ (2521 : 269) กล่าวว่า ความเชื่อมั่น หมายถึง ความคงที่แน่นอนของคะแนน ซึ่งได้จากการวัดนักเรียนกลุ่มเดียวกัน แบบทดสอบฉบับเดียวกันหลายๆ ครั้ง หรือแบบทดสอบสองฉบับที่มีลักษณะเสมอเหมือนกัน หรือภายใต้เงื่อนไขของตัวแปรอื่น ๆ ในการสอบวัด

จากความหมายของความเชื่อมั่นที่บุคคลต่างๆ กล่าวมาผู้วิจัยสรุปได้ว่า ความเชื่อมั่นเป็นความคงที่แน่นอนของคะแนนในการสอบจากผู้สอบกลุ่มเดียวกันทุกครั้ง ด้วยแบบทดสอบฉบับเดียวกันในเวลาที่ต่างกัน

1.2 ทฤษฎีความเชื่อมั่น (Theory of Reliability)

ทฤษฎีการทดสอบแบบมาตรฐานเดิม (Classical Test Theory) เน้นความสำคัญที่ความเชื่อมั่นและความเที่ยงตรงของการวัด และหัวใจสำคัญของความเชื่อมั่นตามทฤษฎีการทดสอบแบบมาตรฐานเดิม คือ มโนคติ (Concept) ของการวัดที่ "คู่ขนาน" กัน (Cronbach and others, 1963 : 137) ซึ่งนำไปสู่การพัฒนาวิธีประมาณค่าความเชื่อมั่นของการวัดทางจิตวิทยาอย่างมาก และได้รับอ้างอิงในหนังสือตำรา และวารสารการวัดผลอย่างแพร่หลาย (Feldt and Brennan, 1989 : 108) การประมาณค่าความเชื่อมั่นของการวัด อาศัยการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ของการวัดหลายๆ ครั้งจากการสอบซ้ำด้วยแบบทดสอบฟอร์มเดียวกัน หรือสองฟอร์มที่คู่ขนานกัน หรือวิเคราะห์ความแปรปรวนของคะแนนสอบภายในฉบับ จากแต่ละส่วนที่สามารถเปรียบเทียบกันได้ (Comparable Parts)(Kristof, 1974:491)

การประมาณค่าความเชื่อมั่นจากการวัดนั้น บุญเชิด ภิญญอนันตพงษ์ (2537 : 1-3) ได้จำแนกตามวิธีการคำนวณได้สามแนวทาง สำหรับสองแนวทางแรกเป็นการประมาณค่าความเชื่อมั่นจากการวัดซ้ำ ถ้าสอบซ้ำด้วยแบบทดสอบฉบับเดียวกัน สัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้จากแบบทดสอบทั้งสองครั้ง เรียกว่าสัมประสิทธิ์ของความคงที่ (Coefficient of Stability) ถ้าสอบซ้ำด้วยแบบทดสอบที่คู่ขนานสองฟอร์ม (Parallel Test Forms) หรือแบบสลับฟอร์ม (Alternate Forms) ซึ่งอาจสอบติดต่อกันทันที หรือสอบทิ้งช่วงเวลาสัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้จากแบบทดสอบทั้งสองฟอร์ม เรียกว่าสัมประสิทธิ์ของความสมมูล (Coefficient of Equivalent) การประมาณค่าความเชื่อมั่นจากการสอบทิ้งช่วงเวลาด้วยฟอร์มที่คู่ขนาน เป็นวิธีที่ให้ค่าประมาณที่ดีที่สุด เพราะสัมประสิทธิ์ชนิดนี้สามารถสะท้อนถึงผลกระทบจากแหล่งความคลาดเคลื่อนในการวัดหมดทุกแหล่ง อย่างไรก็ตามแนวทางทั้งสองดังกล่าวต้องทำการสอบอย่างน้อยสองครั้ง หรือต้องใช้แบบทดสอบอย่างน้อยสองฟอร์ม โดยเฉพาะแนวทางที่สองมักจะไม่สามารถสร้างแบบทดสอบสองฟอร์มให้คู่ขนานกันอย่างแท้จริงได้ จึงไม่เหมาะสมในทางปฏิบัติ สำหรับแนวทางสุดท้ายเป็นวิธีที่หลีกเลี่ยงการสอบซ้ำ โดยอาศัยคะแนนของแบบทดสอบเพียงฉบับเดียวจากการสอบเพียงครั้งเดียว แล้วคำนวณสัมประสิทธิ์ความสอดคล้องภายในแบบทดสอบ (Coefficient of Internal Consistency)

การประมาณค่าความสอดคล้องภายในได้รับความสนใจเป็นอย่างมากจากนักทฤษฎีทางการวัดนับตั้งแต่สเปียร์แมน (Spearman, 1910) และบราวน์ (Brown, 1910) ครอนบัคและคณะ (Cronbach and others, 1963 : 138–139) เริ่มต้นศึกษาเรื่องนี้มาจนถึงปัจจุบัน ได้มีการเสนอสัมประสิทธิ์ในการประมาณค่าความเชื่อมั่นไว้หลายวิธี ซึ่งอาจจัดกลุ่มตามข้อตกลงของระดับความคู่ขนานได้สามรุ่น ดังนี้

รุ่นแรกเป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบที่ แต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิม (Classical Parallel Parts) มีข้อตกลงว่า แต่ละส่วนของแบบทดสอบที่แบ่งต้องมี

- 1) ความเป็นเอกพันธ์ในเนื้อหา หรือวัดคุณลักษณะเดียวกัน
- 2) มีคะแนนจริงเท่ากันและมีความแปรปรวนคลาดเคลื่อนเท่ากัน
- 3) มีคะแนนสอบเฉลี่ยเท่ากัน
- 4) มีความแปรปรวนของคะแนนสอบเท่ากัน
- 5) มีความแปรปรวนร่วมของคะแนนสอบเท่ากัน
- 6) มีความแปรปรวนร่วมของคะแนนสอบกับคะแนนเกณฑ์ภายนอกเท่ากัน

(บุญเชิด ภิญโญอนันตพงษ์, 2537 :2 อ้างอิงมาจาก Gulliksen, 1951:41; Cronbach, 1963 :139 ; Feldt and Brennan, 1989 : 110) จากข้อตกลงดังกล่าวทำให้ไม่สามารถสร้างแบบทดสอบให้แต่ละส่วนมีความคู่ขนานแบบมาตรฐานเดิมได้ใกล้เคียงที่ตกลงไว้ ซึ่งทำให้เกิดการละเมิดข้อตกลงอยู่เสมอ อันเป็นผลให้ไม่สามารถประมาณค่าความเชื่อมั่นได้ถูกต้องแม่นยำ ดังนั้นนักทฤษฎีการทดสอบจึงได้พัฒนาวิธีประมาณค่าความเชื่อมั่นรุ่นสองที่เหมาะสมขึ้นใหม่ เป็นการประมาณค่าความเชื่อมั่นตามแบบจำลองคะแนนจริงสมมูล (Essentially Tau-Equivalent Model)B วิธีนี้ได้ผ่อนปรนเงื่อนไข ข้อ 2) มีคะแนนจริงเท่ากันและมีความแปรปรวนคลาดเคลื่อนเท่ากัน 3) มีคะแนนสอบเฉลี่ยเท่ากันและ 4) มีความแปรปรวนของคะแนนสอบเท่ากัน ให้มีความเป็นไปได้ในทางปฏิบัติมากขึ้น โดยผ่อนปรนให้คะแนนจริงแต่ละส่วนไม่จำเป็นต้องเท่ากันพอดี แต่ยอมให้ต่างกันได้เท่ากับความยากที่ต่างกันในแต่ละส่วนและผ่อนปรนให้แต่ละส่วนให้มีคะแนนสอบเฉลี่ยต่างกันและความแปรปรวนต่างกันเล็กน้อย แต่ยังเป็นไปตามเงื่อนไขข้อ 5) และข้อ 6) แต่ในทางปฏิบัติมีแบบทดสอบบางชนิดอาจต้องแบ่งส่วนให้เหมาะสมตามลักษณะของแบบทดสอบ ทำให้แต่ละส่วนมีขนาดความยาวหรือจำนวนข้อไม่เท่ากัน ซึ่งมีผลกระทบต่อเงื่อนไข 5) มีความแปรปรวนร่วมของคะแนนสอบเท่ากันและ 6) มีความแปรปรวนร่วมของคะแนนสอบกับคะแนนเกณฑ์ภายนอกเท่ากัน แม้แต่ละส่วนประกอบด้วยจำนวนข้อที่เท่ากันก็ตาม แต่เมื่อนำไปสอบกับกลุ่มตัวอย่างแล้ว ปรากฏว่าแต่ละส่วนมีการกระจายของคะแนนมากน้อยต่างกันมาก ๆ แสดงว่าความยาวที่แท้จริงของมัน (Function Lengths) หรือความยาวที่เป็นผลมาจากการสอบ (Functionally) ของแต่ละส่วนมีขนาดไม่เท่ากัน (Feldt 1975 : 558) ดังนั้นจึงไม่สอดคล้องตามแบบจำลองคะแนนสมมูล นักทฤษฎีทางการทดสอบจึงได้นิยามความคู่ขนานขึ้นมากีระดับหนึ่ง เรียกว่า ความคู่ขนานตามแบบจำลองคะแนนจริงสัมพันธ์ (Congeneric Model)

นักทฤษฎีทางการทดสอบได้ทำการศึกษาทดลองและเสนอวิธีประมาณค่าความเชื่อมั่นตามแบบจำลองคะแนนจริงสัมพันธ์ไว้หลายวิธี บุคคลแรกที่ได้เริ่มวางรากฐานทฤษฎีนี้คือ คริสทอฟ (Kristof, 1974 : 491 – 499) ได้เสนอสูตรประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งเป็นสามส่วนย่อยด้วยความยาวขนาดต่างกัน และกล่าวว่าความแปรปรวนคลาดเคลื่อน และค่าความเชื่อมั่นที่คำนวณได้ จะไม่แปรเปลี่ยนไปตามการแบ่งส่วนย่อยของแบบทดสอบ

เฟลด์ต์ (Feldt. 1975 : 557) ได้ปรับปรุงวิธีของคริสทอฟให้สามารถประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งเป็นสองส่วนด้วยความยาวที่ไม่เท่ากัน ต่อมาราชู (Raju. 1977 : 549 – 565) ได้เสนอสูตรประมาณค่าความเชื่อมั่นจากแบบทดสอบที่แบ่งเป็นหลายส่วนย่อยด้วยความยาวขนาดต่าง ๆ ที่ไม่เท่ากัน โดยต้องนำจำนวนข้อมาใช้ในการคำนวณ ส่วนกิลเมอร์และเฟลด์ต์ (Gilmer and Feldt. 1986 : 99 – 111) ได้ปรับปรุงสูตรของคริสทอฟ (Kristof. 1974) ให้สามารถประมาณค่าความเชื่อมั่นจากแบบทดสอบอัตนัย หรือการประเมินบุคคลด้วยคณะกรรมการ ซึ่งมีลักษณะเป็นคะแนนจริงสัมพันธ์หลายส่วน โดยเสนอสูตรการประมาณค่าความเชื่อมั่นสองสูตรคือ r_F และ r_E และเลียว (Liou. 1989 : 153 – 163) ได้เสนอสูตรประมาณค่าความเชื่อมั่นใหม่สองสูตร โดยพัฒนาสูตรจากคริสทอฟ (Kristof. 1974) เป็นสูตร r_{L1} และพัฒนาสูตรของเฟลด์ต์ (Feldt. 1975) เป็นสูตร r_{L2} ให้สามารถประมาณค่าความเชื่อมั่น จากแบบทดสอบที่แบ่งได้หลายส่วน ปรากฏว่าสูตรใหม่ของเลียวให้ผลสอดคล้องกับสูตร r_F และ r_E ของเฟลด์ต์ แต่การคิดคำนวณง่ายกว่า และในปีเดียวกันเฟลด์ต์ (Feldt and Brennan. 1989 : 115) ได้ปรับปรุงสูตรประมาณค่าความเชื่อมั่นของแบบทดสอบที่แบ่งสองส่วน เป็นสูตรที่สามารถประมาณค่าจากแบบทดสอบที่แบ่งเป็นหลายส่วน โดยมีได้แสดงที่มาและผลการทดลอง

วิลเลียม(William, MK2001) นำเสนอทฤษฎีของการประมาณค่าความเชื่อมั่นมีทั้งหมด 4 กลุ่ม ดังนี้

1. ความเชื่อมั่นระหว่างผู้สังเกตหรือผู้ประเมิน (Inter-Rater or Inter-Observer Reliability) เป็นผลจากการประเมินระดับของความแตกต่างจากการตรวจ หรือการสังเกต ที่มีความคงที่ และใกล้เคียงกันมากที่สุด
2. ความเชื่อมั่นจากการสอบซ้ำ (Test-Retest Reliability) เป็นผลที่เกิดจากการประเมินที่มีผลคงที่ เมื่อมีการวัดซ้ำจากเครื่องมือวัดเดียวกัน
3. ความเชื่อมั่นจากแบบทดสอบคู่ขนาน (Parallel-Forms Reliability) ผลจากการประเมินที่มีผลคงที่เมื่อทดสอบด้วยแบบทดสอบคู่ขนานกัน
4. ความเชื่อมั่นจากความคงที่ภายใน (Internal Consistency Reliability) ผลจากการประเมินที่มีค่าคงที่ภายในด้วยแบบทดสอบฉบับเดียวกัน

ผู้วิจัยสนใจศึกษาเฉพาะความเชื่อมั่นระหว่างผู้สังเกตหรือผู้ประเมิน (Inter-Rater or Inter-Observer Reliability) มีรายละเอียดที่น่าสนใจดังนี้ ความเชื่อมั่นระหว่างผู้ประเมินเป็นหนึ่งในสี่ของการประมาณค่าความเชื่อมั่นที่ใช้ประเมินระดับความแตกต่างของผู้ประเมินที่จะมีความคงที่ในการประเมินสิ่งเดียวกัน ซึ่งการนำมนุษย์มาเป็นส่วนหนึ่งของการวัดและประเมินผล ซึ่งมีความกังวลใจเกี่ยวกับผลที่จะได้รับมีความเชื่อถือได้หรือไม่ วิธีใดที่จะทำให้ทราบผลการประเมินของผู้ประเมิน 2 คน มีค่าคงที่ได้ ซึ่งวิลเลียมเสนอไว้มี 2 วิธีสำคัญ ๆ คือ 1) นำผลการวัดของผู้ประเมินตรวจผลงานจากเครื่องมือการสังเกตแล้ว เราสามารถคำนวณโดยหาร้อยละของความสอดคล้องระหว่างผู้ประเมิน 2) อีกวิธีสำหรับการประมาณค่าความเชื่อมั่นที่เหมาะสม เมื่อมีผลการวัดที่ต่อเนื่องอยู่ เราสามารถคำนวณหาสหสัมพันธ์ระหว่างการประเมินของผู้ประเมิน 2 คน

1.3 ความเชื่อมั่นของผู้ประเมิน (Rater Reliability)

จากการศึกษาเอกสารพบว่า มีนักวัดผลการศึกษาทั้งในและต่างประเทศได้กล่าวถึง ความหมายและความสำคัญของความเชื่อมั่นของผู้ประเมินให้คะแนนดังนี้

ฮอปกินส์ และ แอนเตส (Hopkins and Antes. 1990 : 550) ให้ความหมายของความเชื่อมั่นของผู้ประเมินคะแนนที่มีตั้งแต่สองคนขึ้นไป จะมีความเห็นสอดคล้องกันเกี่ยวกับจุดเน้นหรือจุดสำคัญต่าง ๆ ของคำตอบที่ผู้ประเมินคะแนนตามความเห็นของผู้ประเมินแต่ละคน ค่าความเชื่อมั่นของผู้ประเมินสูง ก็หมายถึงค่าสัมประสิทธิ์สหสัมพันธ์ที่สูงตามไปด้วย ซึ่งผู้ประเมินข้อสอบที่ดีได้ ควรเป็นผู้เขียนข้อสอบหรือผู้สอบวิชานั้น ๆ จริง ๆ (Hopkins and Antes, 1990 : 235)

อุเบอแซก (Uebersax. 2000) ได้เสนอดัชนีความสอดคล้องจากคะแนนดิบ (Raw Agreement Indices) ไว้ 3 กลุ่ม คือ

1. ผู้ประเมิน 2 คน ประเมินให้คะแนนแบบ 2 ขั้ว (Two Rater, Dichotomous Rating)
2. ผู้ประเมิน 2 คน ประเมินให้คะแนนแบบหลายขั้ว (Two Rater, Polytomous Ratings)
3. กรณีทั่วไป (Generalized Case)

ในที่นี้ผู้วิจัยสนใจศึกษาวิธีการของ กลุ่มที่ 1 และกลุ่มที่ 2

1. ผู้ประเมิน 2 คน ประเมินให้คะแนน 2 ขั้ว (Two Rater, Dichotomous Rating)

การตัดสินใจคะแนนจากผู้ประเมินจำนวน 2 คน (หรือ ผู้เชี่ยวชาญ , ผู้ตัดสิน, ผู้วินิจฉัย หรือ อื่นๆ) มารวมกัน ดังตารางต่อไปนี้

ตาราง 1 ผลรวมของการประเมินให้คะแนนแบบ 2 ขั้ว จากผู้ประเมิน 2 คน

ผู้ประเมินคนที่ 1	ผู้ประเมินคนที่ 2		รวม
	-	+	
-	a	b	a + b
+	c	d	c + d
รวม	a + c	b + d	N

ค่า a, b, c และ d ในที่นี้เป็นผลรวมจากความคิดเห็นที่สอดคล้องกันทั้งหมดของผู้ประเมินแต่ละคนคือ ผู้ประเมินคนที่ 1 และคนที่ 2 แบ่งได้ 2 ลักษณะ คือ

สัดส่วนของความคิดเห็นที่สอดคล้องกันทั้งหมด (Proportion of Overall Agreement)

สัดส่วนของคะแนนจากความคิดเห็นที่สอดคล้องกันทั้งหมด กำหนด p_0 เป็นสัดส่วนของกรณีที่มีผู้ประเมินคนที่ 1 และ 2 ดังนี้

$$p_0 = \frac{a + d}{a + b + c + d} = \frac{a + d}{N} \dots\dots\dots (1)$$

ผลการคำนวณ ถ้าค่า p_0 สูง คือมากกว่า .99 ก็สรุปได้ว่าความสอดคล้องระหว่างความคิดเห็นของผู้ประเมินไม่แตกต่างกัน

สัดส่วนของความคิดเห็นที่เฉพาะเจาะจง (Proportion of specific Agreement)

สัดส่วนของความคิดเห็นทางบวก ($p_s +$) และความคิดเห็นทางลบ ($p_s -$) จะคำนวณได้จากสูตร ดังนี้

$$p_s + = \frac{2a}{2a + b + c}; p_s - = \frac{2a}{2d + b + c} \dots\dots\dots (2)$$

สัดส่วนนี้แปลความหมายได้คล้ายกับการประมาณความเป็นไปได้ที่ตั้งเงื่อนไขไว้ ตัวอย่าง คือ $p_s +$ เป็นค่าประมาณจากผู้ประเมินที่ให้คะแนนทางบวกได้คล้าย ๆ กับผู้ประเมินคนอื่น

2. ผู้ประเมิน 2 คน ประเมินให้คะแนนแบบหลายชั่ว (Two Raters, Polytomous Rating)

ใช้การตัดสินจากผู้ประเมิน 2 คน ที่ให้คะแนนแบบหลายชั่ว โดยให้ C เป็นจำนวนของกลุ่มหรือระดับในการให้คะแนน ผลของการให้คะแนนจากผู้ประเมิน 2 คน จะถูกรวบรวมตามวิธี C x C ดังตารางต่อไปนี้

ตาราง 2 ผลรวมของการให้คะแนนแบบหลายชั่ว โดยผู้ประเมิน 2 คน

ผู้ประเมิน คนที่ 1	ผู้ประเมินคนที่ 2			รวม
	1	2	...C	
1	n_{11}	n_{12}	n_{1c}	$n_{1.}$
2	n_{21}	n_{22}	n_{2c}	$n_{2.}$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
C	n_{c1}	n_{c2}	n_{c3}	$n_{c.}$
รวม	$n_{.1}$	$n_{.2}$	$n_{.c}$	N

ให้ n_{ij} แทน จำนวนของกรณีที่ได้วางแผนให้คะแนนชุด i จากผู้ประเมินคนที่ 1 และชุด j โดยผู้ประเมินคนที่ j เมื่อ $i, j = 1, \dots, C$. แทน ค่าผลรวม (magical Sum) เช่น $n_{.j}$ เป็นผลรวมของ n_{ij} ที่ $j = 1, \dots, C$ แบ่งได้ 2 ลักษณะ คือสัดส่วนของความคิดเห็นที่สอดคล้องกันทั้งหมด (Proportion of Overall Agreement)

การคำนวณ p_0 ซึ่งก็คือผลรวมของความถี่จากการตรวจให้คะแนน ซึ่งคำนวณด้วยสูตร ดังนี้

$$p_0 = \frac{1}{N} \sum_{i=1}^C n_{ii} \dots\dots\dots (3)$$

สัดส่วนของความคิดเห็นที่แตกต่าง (Proportion of Specific Agreement)

ใช้คำตอบจากตาราง 2 สัดส่วนของความคิดเห็นที่แตกต่างกัน จากข้อมูลชุด i ใช้สูตรดังนี้

$$p_s(i) = \frac{1n_{ii}}{n_{i.} + n_{.i}} \dots\dots\dots (4)$$

เจมส์และคนอื่น ๆ (James and others. 1984 ; 1993) ได้ทำการพัฒนาสัมประสิทธิ์ความเชื่อมั่นระหว่างผู้ประเมิน (Interrater reliability Coefficient (r_{wg}) ซึ่งจะประเมินความสอดคล้องระหว่างผู้ประเมิน โดยกำหนดให้ r_{wg} เป็นการลดสัดส่วนความแปรปรวนที่ผิดพลาดสำหรับการให้คะแนนที่ได้มา เพื่อเทียบกับการจัดแบ่งที่แสดงถึงการตอบแบบสุ่ม (ตัวอย่าง เช่น การแบ่งในกรณีแต่ละคะแนนในเกณฑ์ของอันดับคะแนนมีความเป็นไปได้เท่า ๆ กันในการนำไปใช้) กล่าวคือ

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

โดยที่ S_x^2 เท่ากับค่าความแปรปรวนของผลการประเมินให้คะแนน และ σ_E^2 เท่ากับค่าความแปรปรวนของประชากร ในการจัดแบ่งอันดับที่ได้จากการสุ่ม (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 หรือหากผลการประเมินให้คะแนนเหมือนกัน (เช่น $S_x^2 = 0$) ค่า $r_{wg} = 1$ ซึ่งชี้ให้เห็นว่ามีความสอดคล้องโดยสมบูรณ์ระหว่างผู้ประเมินทั้งหมด

ค่า r_{wg} ที่ได้ควรจะเป็นดัชนีบ่งบอกความสอดคล้อง (Index of Agreement) ไม่ใช่ค่าความคงที่ (Consistency) เจมส์ (James. 1993) อธิบายว่า การให้คะแนนบางประเภท อาจมีระดับความคงที่สูง (เช่น ผู้ตรวจ 2 คน หรือมากกว่านั้นให้คะแนนออกมาคล้ายคลึงกัน เมื่อให้คะแนนตัวอย่างหลาย ๆ ตัวอย่างบนมาตรฐานเดียวกัน) แต่ค่าการให้คะแนนที่แท้จริงที่ให้ผู้ประเมินคนเดียวอาจจะต่างกันไปมาก ให้พิจารณาถึงกรณีที่ผู้ประเมินคนที่ 1 ให้คะแนน "4" ทุกครั้ง ผู้ประเมินคนที่ 2 ให้คะแนน "2" และทุกครั้งที่ผู้ประเมินคนที่ 1 ให้คะแนน "3" ผู้ประเมินคนที่ 2 ให้คะแนน "1" ค่าสหสัมพันธ์ระหว่างความคงที่ระหว่างผู้ประเมิน 2 คน นั้นสูง มาก แต่ระดับความสอดคล้องระหว่างผู้ประเมินนี้จะต่ำเพราะว่าพวกเขาไม่ให้ค่าที่เป็นคะแนนซ้ำกันเหมือนเดิม ข้อดีของสัมประสิทธิ์ r_{wg} ก็คือ สัมประสิทธิ์ที่ได้จะแปรผันได้ตามค่าสมบรูณ์ในการให้คะแนน เช่นเดียวกับการจัดอันดับคะแนนการประเมินให้คะแนน

1.4 สูตรที่ใช้คำนวณหาความเชื่อมั่นของการประเมินผลงาน

สูตรที่ใช้คำนวณหาความเชื่อมั่นของการประเมินในครั้งนี้ ผู้วิจัยได้สูตรของแฮมเบิลตันและโนวิก สูตรแคปปาของสก็อต และสูตร RAI ของเบอร์รี – สต็อกค์ และคณะ

1.4.1 สูตรของแฮมเบิลตันและโนวิก

แฮมเบิลตัน และโนวิก (Subkoviak 1976 : 266 citing Hambleton and Novick. 1973) ได้เสนอแนะว่า ความเชื่อมั่นของการตัดสินเพื่อจำแนกผู้รอบรู้ อาศัยนิยามในรูปของความสอดคล้องในการตัดสินความรอบรู้จากการสอบด้วยแบบทดสอบฉบับเดียวกันสองครั้ง หรือจากการสอบด้วยแบบทดสอบคู่ขนานสองฉบับ แล้วนำผลการสอบไปแจกแจง และหาค่าสัดส่วนความสอดคล้องในการตัดสินความรอบรู้จากผลการสอบทั้งสองฉบับหรือสองครั้ง ดังนี้

ฟอร์ม ข. (หรือสอบครั้งที่ 2)

ฟอร์ม ก (สอบครั้งที่ 1)	ผ่าน ไม่ผ่าน	ผ่าน	ไม่ผ่าน
		P_{11}	P_{22}

สัดส่วนของความสอดคล้องในการตัดสินใจคือ สัดส่วนของการตัดสินใจว่ารอบรู้ตรงกันทั้งสองฉบับหรือสองครั้ง และสัดส่วนของการตัดสินใจว่าไม่รอบรู้กันทั้งสองฉบับหรือสองครั้งซึ่งสามารถเขียนเป็นสูตรได้ดังนี้

$$P_0 = \sum_{i=1}^2 P_{ii} = P_{11} + P_{22}$$

P_0 แทน สัดส่วนของความสอดคล้องในการประเมินเพื่อจำแนก

P_{11} แทน สัดส่วนของผู้ประเมินว่าประเมินตรงกันทั้งสองครั้ง

P_{22} แทน สัดส่วนของผู้ประเมินว่าประเมินไม่ตรงกันทั้งสองครั้ง

1.4.2 สูตรแคลปา ของสก็อต

โอเบอร์ ริชาร์ดและคณะ (Ober, Richard L. Bentley, Ernest L., and Miller. 1971 : 80) ได้ศึกษาความเชื่อมั่นภายในและระหว่างกลุ่มของสก็อต (Scott's) โดยเป็นวิธีการดังกล่าวเป็นวิธีที่รวดเร็วในการหาสัมประสิทธิ์ของสก็อต การวิเคราะห์ในการหาความเชื่อมั่นภายในและระหว่างกลุ่ม โดยใช้สูตรของสก็อต เป็นดังนี้

$$\text{ความเชื่อมั่น (Reliability)} = \frac{P_0 - P_e}{1.00 - P_e}$$

P_0 แทน ความสอดคล้องระหว่างผู้ประเมิน 5 คน

P_e แทน ความสอดคล้องระหว่างผู้ประเมิน 5 คนที่เกิดขึ้นในตัวอย่างโดยความบังเอิญ

1.4.3 สูตร RAI ของเบอร์รี – สต็อกค์

เบอร์รี – สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia Laurie, Brad S. Chissom. 1996) ได้ศึกษาดัชนีความสอดคล้องระหว่างผู้ตรวจให้คะแนน ได้ผลการศึกษาที่สามารถสรุปวิธีการคำนวณได้จากผลการสังเกตพฤติกรรมในระดับช่วงคะแนน (I) ที่มีของดัชนีจากค่ามากที่สุด คือ 1 ไปหาค่าน้อยที่สุด คือ 0 สรุปแนวคิดจากการตรวจให้คะแนนของผู้ตรวจหลาย ๆ คนได้ 6 วิธี ดังนี้

1. RAI สำหรับข้อสอบ 1 ข้อ ผู้ตรวจให้คะแนน 2 คน, สำหรับพฤติกรรมเดียว เป็นสูตรคำนวณจากผลการให้คะแนนของผู้ตรวจให้คะแนน 2 คน ที่ให้คะแนนในพฤติกรรมเดียวกัน ที่มีพิสัยของดัชนีจากค่ามากที่สุด (I) ถ้าผู้ตรวจให้คะแนน 2 คน ให้คะแนนเหมือนกันค่า RAI = 1 ถ้าผู้ตรวจให้คะแนน 2 คน ให้คะแนนตรงกันข้าม ค่า RAI = 0

2. RAI ที่ใช้ผลรวมของคะแนนที่มีผู้ตรวจให้คะแนน 2 คน กรณีที่ประเมิน K พฤติกรรม ผู้ประเมิน 2 คน และกลุ่มตัวอย่างเพียงคนเดียว สูตรนี้สามารถแสดงค่า RAI สำหรับ K พฤติกรรม
3. RAI ที่ใช้ผลรวมของคะแนนที่มีผู้ตรวจให้คะแนน 2 คน กลุ่มตัวอย่าง N คน กรณีที่ประเมิน K พฤติกรรม
4. RAI สำหรับกรณีที่ประเมินพฤติกรรมเดียว ผู้ประเมิน M คน และกลุ่มตัวอย่างเพียงคนเดียว
5. RAI สำหรับกรณีที่ประเมิน K พฤติกรรม ผู้ประเมิน M คน และกลุ่มตัวอย่างเพียงคนเดียว
6. RAI สำหรับกรณีที่ประเมิน K พฤติกรรม ผู้ประเมิน M คน และกลุ่มตัวอย่าง N คน

จากข้อมูลการศึกษาค้นคว้าเอกสารทั้งหมดนี้ ผู้วิจัยได้นำสูตรที่ 4 มาใช้ในการวิเคราะห์ ดังมีรายละเอียดของสูตร ดังนี้

$$RAI = 1 - \frac{\sum |R_m - \bar{R}|}{(M-1)(I-1)} \dots\dots\dots (4)$$

เมื่อ	R	แทน	ผลการประเมินของผู้ประเมินให้คะแนน
	M	แทน	จำนวนผู้ประเมินให้คะแนนทั้งหมด
	I	แทน	จำนวนช่วงคะแนน (1,2,3,4)

2. การวัดภาคปฏิบัติ

2.1 ความหมายของการวัดภาคปฏิบัติ

นักการศึกษาหลายท่านได้ให้ความหมายของการวัดภาคปฏิบัติ ก่อนข้างจะมีความหมายในทำนองเดียวกันดังต่อไปนี้

อาห์แมน และ กล็อก (Ahmann and Glock. 1967 : 186) ได้ให้ความหมายของแบบทดสอบวัดภาคปฏิบัติไว้ว่าเป็นแบบทดสอบเพื่อพิจารณาการกระทำในการประเมินผลวิธีการปฏิบัติงานและผลผลิตของงาน

เมห์เรนส์ และเลห์แมน (Mehrens and Lehmann. 1984 : 206) ได้ให้ความหมายของแบบทดสอบวัดภาคปฏิบัติไว้ว่า แบบทดสอบวัดภาคปฏิบัติโดยทั่วไปจะเป็นแบบทดสอบ 1 ใน 3 ของประเภทต่อไปนี้

1. แบบทดสอบภายใต้สภาวะการณ์จำลองขึ้นมา (Tests Under Simulated Conditions) เช่น การฝึกหัด (Training) ของนักบินภายใต้อุปกรณ์ฝึกภาคสนามสำหรับนักบิน (Link Trainer) เป็นตัวอย่างของแบบทดสอบชนิดนี้ ข้อจำกัดของวิธีดำเนินการของแบบทดสอบชนิดนี้ก็คือ พฤติกรรมในสถานการณ์จำลองขึ้นมามีที่แตกต่างกันอย่างชัดเจน เมื่อเปรียบเทียบกับสถานการณ์ที่เป็นจริง ดังนั้นผลที่ได้จากการประเมินการปฏิบัติจึงมีความถูกต้องน้อย

2. แบบทดสอบตัวอย่างงาน (Work Sample Tests) เป็นแบบทดสอบที่เชื่อถือได้ (Reliable) และสมเหตุสมผล (Valid) มากที่สุด เพราะว่าผู้เข้าสอบจะทำการผลิตบางสิ่งบางอย่างแท้จริง

3. แบบทดสอบการจำได้ (Recognition Tests) เป็นแบบทดสอบที่จะวัดความสามารถของผู้ที่เข้าสอบเกี่ยวกับความสามารถในการจำลักษณะเฉพาะ (Characteristics) ของผลผลิต (Products) และการปฏิบัติได้หรือไม่ หรือวัดความสามารถในการจำแนกแยะแยก (Identify) เช่นตัวอย่างเกี่ยวกับพืชบางชนิด

ธอร์นไคค์ (Thomdike. 1960 : 238) ได้กล่าวว่า แบบทดสอบวัดภาคปฏิบัติบางครั้งก็ถูกนำมาใช้ให้ ความหมายของแบบทดสอบอาชีพ (Trade Test) เพื่อใช้ประเมินความสามารถในทางอาชีพของช่างฝีมือ (Skilled Worker)

มาร์แชล และลอยด์ (Marshall and Loyde. 1971 : 135) ได้ให้ความหมายของแบบทดสอบ วัดภาคปฏิบัติไว้ว่า เป็นแบบทดสอบที่เกี่ยวกับการตอบสนองที่เป็นการกระทำของผู้ถูกทดสอบ โดยปกติ แล้วการทดสอบจะเกิดขึ้นได้ต้องจัดให้ผู้ถูกทดสอบได้อยู่ในสถานการณ์ที่เป็นจริงหรือคล้ายของจริงให้มากที่สุด นอกจากนี้ยังสามารถจัดรูปแบบทดสอบประเภทนี้ไว้ 3 ความหมาย คือ

1. แบบทดสอบวัดภาคปฏิบัติที่เกี่ยวกับความสามารถทางสมอง ด้านความคิด ส่วนใหญ่จะเกี่ยวข้องกับ การสอบทักษะด้านภาษา การฟัง การพูด และการกระทำ

2. แบบทดสอบวัดภาคปฏิบัติที่ทดสอบความสามารถในการใช้เครื่องจักรและเครื่องมือต่าง ๆ ประกอบในการทำงานสิ่งใดสิ่งหนึ่งให้ประสบความสำเร็จ

3. แบบทดสอบวัดภาคปฏิบัติในความหมายนี้ อาจรวมไปถึงการใช้การเขียนตอบบนกระดาษได้ด้วย

สแตนลีย์ (Stanley. 1975 : 186) ได้ให้ความหมายของแบบทดสอบภาคปฏิบัติไว้ว่าเป็นแบบทดสอบ เพื่อพิจารณาการกระทำในการประเมินผลวิธีปฏิบัติงานและผลผลิตของงาน

ไพศาล หวังพานิช (2526 : 89) ได้ให้ความหมายไว้ว่า การวัดผลภาคปฏิบัติเป็นการวัดความ สามารถในการปฏิบัติ ซึ่งเป็นการวัดที่ให้ผู้เรียนได้แสดงพฤติกรรมตรงออกมา ด้วยการกระทำโดยถือว่าการ ปฏิบัติเป็นความสามารถในการผสมผสานหลักวิธีการต่าง ๆ ที่ได้รับการฝึกฝนมาให้ปรากฏออกมาเป็น ทักษะของผู้เรียน

สมบูรณ์ ชิตพงษ์ (2522 : 18) ได้ให้ความหมายของการวัดภาคปฏิบัติว่าเป็นการวัดที่มีจุดประสงค์ เพื่อต้องการให้ผู้สอบได้ปฏิบัติให้ดู การทดสอบแบบนี้ต้องการวัดวิธีการหรือดูผลงานในการปฏิบัติ

เชิดศักดิ์ ไชวสินธุ์ (2529 : 16) ได้ให้ความหมายของการวัดภาคปฏิบัติไว้ว่าเป็นเครื่องมือที่ ออกแบบเพื่อวิเคราะห์และวัดทักษะของนักเรียนในด้านการปฏิบัติหรือการกระทำที่ให้เลือกปฏิบัติภายใต้เงื่อนไข ที่ได้ควบคุมไว้อย่างดี

เผียน ไชยศรี (2529 : 37) กล่าวว่า การวัดผลงานภาคปฏิบัติ เป็นการวัดความสามารถของ บุคคลในการทำงานอย่างใดอย่างหนึ่ง โดยบุคคลนั้นได้ลงมือปฏิบัติการจัดกระทำ (Manipulate) ซึ่งมีการ เกี่ยวข้องหรือสัมพันธ์กับสิ่งที่อยู่ในลักษณะของรูปธรรม (Materials or Physical Objects) โดยทางกาย หรือการรับรู้ทางประสาทสัมผัส

ส.วาสนา ประวาลพฤษ์ (2533 : 8) ได้ให้ความหมายของการวัดผลภาคปฏิบัติในแง่ของการวัด ผลสัมฤทธิ์ทางการเรียน หมายถึง การให้ผู้สอบทำงานในกลุ่มตัวอย่างของงานที่จำเป็นในงานหนึ่ง โดยจะ กำหนดลักษณะของงานหรือเครื่องมือ หรือผลผลิตที่จะวัดซึ่งจะวัดในรูปของทักษะในการดำเนินงาน

นิโลบล นิมกักรัตน์ (2531 : 24) ได้ให้ความหมายของการวัดภาคปฏิบัติว่า เป็นการวัดความสามารถ ของบุคคลในการทำงานอย่างใดอย่างหนึ่ง อาจจะเริ่มต้นวัดตั้งแต่ ขั้นเตรียม ขั้นลงมือปฏิบัติ ขั้นผลงาน ทั้งนี้ขึ้นอยู่กับวัตถุประสงค์ในการวัดแต่ละครั้ง

สุภาพ ฉัตรภรณ์ (2539 : 152) กล่าวว่า การวัดผลภาคปฏิบัติ เป็นการวัดผลความรู้ความสามารถ ในการทำงานของผู้เรียนจากสิ่งต่างๆ ดังนั้น ทักษะในการทำงานตามขั้นตอน (Skill) ความถูกต้องแม่นยำ (Accuracy) และผลงาน (Product)

สมศักดิ์ สินธุระเวชญ์ (2530 : 127-130) กล่าวว่า การวัดการปฏิบัติงานจริงเป็นการวัดผลการปฏิบัติงานจริงซึ่งจะใช้วิธีการสังเกต เพราะการสังเกตทำให้สามารถเรียนรู้เรื่องราวของนักเรียนแต่ละคนได้ แต่การสังเกตที่ไม่ได้มีการเตรียมรายละเอียดต่าง ๆ หรือใช้วิธีการที่ดี ก็จะทำให้ขาดความเชื่อมั่นได้

จากความหมายของการวัดผลภาคปฏิบัติดังกล่าว พอสรุปได้ว่า การวัดผลภาคปฏิบัติเป็นการวัดความสามารถของบุคคลในด้านทักษะ ความคิดสร้างสรรค์ ความรู้ ในการลงมือกระทำต่อผลงานตามจุดมุ่งหมายที่ได้กำหนดขึ้น

2.2 พฤติกรรมการเรียนรู้ทางภาคปฏิบัติ

นักการศึกษาหลายท่านได้แบ่งระดับพฤติกรรมทางภาคปฏิบัติเป็นขั้น ๆ แตกต่างกันไป ดังนี้

ซิมป์สัน (Simpson, 1966) ได้แบ่งระดับของพฤติกรรมทางด้านการปฏิบัติ โดยเรียงจากการรับรู้ต่ำสุด เป็นไปตามลำดับขั้นดังนี้

1. การรับรู้ (Perception) เป็นขั้นสำคัญขั้นแรก ในการปฏิบัติหรือลงมือทำกิจกรรมเป็นกระบวนการของการรับรู้ เกี่ยวกับวัตถุ คุณภาพ หรือความสัมพันธ์ โดยอาศัยประสาทสัมผัส ขั้นนี้เป็นพื้นฐานของวงจรที่เกี่ยวกับสถานการณ์การแปลผลการกระทำ ซึ่งนำไปสู่กิจกรรม ทักษะ การรับรู้ที่จัดอยู่ในขั้นนี้แบ่งเป็น 3 ระดับขั้นของกระบวนการรับรู้

1.1 การเร้าความรู้สึก (Sensory Stimulation) การกระตุ้นของสิ่งเร้าต่อประสาทความรู้สึกหนึ่งอย่างหรือมากกว่าหนึ่งอย่าง เช่น ทางหู (Auditory) การได้ยินหรือประสาทที่เกี่ยวกับการได้ยิน ทางตา (Visual) เกี่ยวกับภาพพจน์ในสมอง หรือภาพพจน์จากการมองดู ทางมือ (Tactile) เป็นการรับรู้ความรู้สึกได้จากการสัมผัส ทางลิ้น (Taste) เป็นการรับรู้ความรู้สึกโดยอาศัยรส หรือการนำเข้าปาก ทางดมกลิ่น (Smell) รับรู้โดยการได้รับกลิ่นจากจมูก ทางกล้ามเนื้อ (Kinesthetic) ความรู้สึกทางกล้ามเนื้อเกี่ยวกับความรู้สึกจากกิจกรรมของเครื่องรับ ซึ่งอยู่ในกล้ามเนื้อเอ็นและข้อต่อ

1.2 การเลือกแนวทางปฏิบัติ (Cue Selection) การตัดสินใจที่จะเลือกกิจกรรมที่คนจะสนองตอบเพื่อให้สอดคล้องกับความต้องการในการปฏิบัติงาน

1.3 การแปลความหมาย (Translation) เกี่ยวกับการรับรู้ต่อการปฏิบัติกิจกรรมที่เป็นกระบวนการทางสมองในการตัดสินใจตีความหมายของการกำหนดพฤติกรรมในการกระทำ การแปลเกี่ยวข้องกับสัญลักษณ์ คือการมีภาพพจน์หรือการย้อนรำลึกถึงบางอย่าง “การมีความคิด” อันเป็นผลของการกำหนดพฤติกรรมที่ได้รับมา

2. ความพร้อมปฏิบัติ (Set) ความพร้อมเป็นการปรับตัวให้เตรียมพร้อมต่อการกระทำบางอย่างหรือประสบการณ์บางอย่าง ความพร้อมมี 3 ลักษณะ คือ

2.1 ความพร้อมทางสมอง (Mental Set) คือความพร้อมในเชิงความคิด ที่จะกระทำกิจกรรมทางทักษะบางอย่าง ความพร้อมทางสมองเป็นสิ่งที่ต้องมีมาก่อนและเกี่ยวข้องถึงระดับของการรับรู้

2.2 ความพร้อมทางร่างกาย (Physical Set) คือความพร้อมในลักษณะที่มีการปรับสภาพร่างกายซึ่งจำเป็นสำหรับการกระทำบางอย่าง ความพร้อมทางร่างกายเกี่ยวข้องกับความพร้อมด้านประสาท การรับรู้ต่าง ๆ คือการรับรู้ความรู้สึก หรือการมุ่งไปยังความตั้งใจของประสาทสัมผัส และท่าทางที่จำเป็น หรือการจัดท่าทางตำแหน่งของร่างกาย

2.3 ความพร้อมทางอารมณ์ (Emotional Set) คือความพร้อมในรูปของเจตคติที่ดีที่พึงปรารถนา ต่อกิจกรรมปฏิบัติที่เกิดขึ้น นั่นคือความตั้งใจตอบสนอง

3. การตอบสนองตามแนวทางที่กำหนดให้ (Guided Response) เป็นขั้นต้นในการพัฒนาทักษะ การเน้นอยู่ที่ความสามารถซึ่งเป็นส่วนสำคัญของทักษะที่ซับซ้อนยิ่งขึ้น การตอบสนองตามแนวทางนี้เป็นการ แสดงออกของแต่ละคน ภายใต้การแนะนำแนวทางของผู้สอน จำแนกได้ 2 อย่าง คือการเลียนแบบและการ ลองผิดลองถูก

3.1 การเลียนแบบ (Imitation) หมายถึง การดำเนินการในลักษณะที่เป็นการตอบสนองโดยตรง ตามการรับรู้ของบุคคลที่กระทำการนั้น ๆ

3.2 การลองผิดลองถูก (Trial and Error) เป็นความพยายามที่จะตอบสนองแบบต่าง ๆ ตาม ปกติจะมีหลักการซ่อนเร้นอยู่ภายใต้การตอบสนองแต่ละอย่างจนกระทั่งประสบผลสำเร็จ การตอบสนองที่ เหมาะสมเป็นเรื่องที่ต้องให้ตอบสนองตามความต้องการของงานที่ปฏิบัติ นั่นคือ “ทำงานให้เสร็จ” หรือทำงาน อย่างมีประสิทธิภาพ ซึ่งการตอบสนองที่เหมาะสมได้รับการคัดเลือกจากพฤติกรรมต่าง ๆ อาจอาศัยอิทธิพล ของรางวัลหรือการลงโทษ

4. ขั้นทักษะ (Mechanism) คือการตอบสนองที่เกิดการเรียนรู้จนกลายเป็นนิสัย ผู้เรียนมีความเชื่อมั่น และมีความชำนาญมากพอที่จะปฏิบัติงานนั้นๆ กิจกรรมเป็นส่วนหนึ่งของสิ่งที่จะทำได้ ซึ่งสามารถจะตอบ สอนสิ่งเร้าและข้อเรียกร้องของสถานการณ์ซึ่งต้องการการตอบสนองอย่างเหมาะสม การตอบสนองจะซับซ้อน กว่าระดับที่ผ่านมา และเกี่ยวข้องกับรูปแบบอย่างที่ใช้กระทำกิจกรรมนั้น ๆ

5. การปฏิบัติงานที่ยากและซับซ้อน (Complex Overt Response) การตอบสนองในขั้นนี้ต้อง อาศัยทักษะขั้นสูง โดยตอบสนองด้วยความราบรื่นมีประสิทธิภาพสูง และใช้เวลาน้อยที่สุด แบ่งออกเป็น 2 แบบ คือ

5.1 การตอบสนองโดยไม่ลังเล (Resolution of Uncertainty) เป็นการปฏิบัติโดยปราศจากความ ลังเลใจ โดยสมองจะรู้ภาพขั้นตอนที่ต้องการด้วยความมั่นใจ สามารถทำในสิ่งที่ซับซ้อนได้

5.2 การตอบสนองแบบอัตโนมัติ (Automatic Performance) ในระดับนี้แต่ละคนสามารถใช้ทักษะ ของกล้ามเนื้อ ซึ่งประสานกันอย่างดีโดยสะดวก รวมทั้งควบคุมกล้ามเนื้อได้

6. การดัดแปลงให้เหมาะสม (Adaptation) เป็นการเปลี่ยนกิจกรรมในทางเคลื่อนไหวโดยสมองให้ เคลื่อนไหวได้เหมาะสมกับความต้องการหรือสถานการณ์ของปัญหา

7. การริเริ่ม (Origination) หมายถึง การริเริ่มรูปแบบการเคลื่อนไหวใหม่ ๆ ที่เหมาะสมกับ สถานการณ์เฉพาะอย่างหรือปัญหาเฉพาะเรื่อง ผลการเรียนรู้ระดับนี้เป็นการพัฒนาทักษะขั้นสูง

เดฟ (โกวิท ประวาลพุกษ์. 2523 : 28 ; อ้างอิงมาจาก Dave. 1969) ได้แบ่งพฤติกรรมด้าน การปฏิบัติโดยเรียงตามระดับการประสานกัน คือ

1. การเลียนแบบ (Imitation) ทำเลียนแบบซ้ำ ๆ โดยที่ยังไม่ได้ผลสมบูรณ์
2. ยักย้ายถ่ายเท (Manipulation) ทำตามแบบโดยมีคำสั่งชี้แจงที่จะพัฒนาทักษะ
3. ประณีต (Precision) ทำอย่างมีทักษะโดยปราศจากคำแนะนำหรือรูปแบบ
4. มีศิลปะ (Articulation) ทำต่อเนื่องประสานกัน เพื่อความถูกต้องและควบคุมได้
5. ทำได้อย่างธรรมชาติ (Naturalization) มีความสามารถระดับสูง ทำอย่างอัตโนมัติ

ฮาร์โรว์ (โกวิท ประวาลพุกษ์ และสมศักดิ์ สินธุระเวชญ์. 2523 : 29 ; อ้างอิงจาก Harrow. 1972) แบ่งพฤติกรรมด้านการปฏิบัติตามระดับพฤติกรรมการเคลื่อนไหวดังนี้

1. การเคลื่อนไหวตอบโต้ (Reflex Movement) โดยการเคลื่อนไหวหรือไม่เคลื่อนไหวหรือผสมผสานกัน
 2. การรับรู้ (Perceptual) โดยการรู้สึกทางการเคลื่อนไหวทางตา ทางหู ทางการสัมผัส และทางการประสานกันของอวัยวะดังกล่าวมาแล้ว
 3. ความสามารถทางกาย (Physical Ability) หมายถึง ความทนทาน ความแข็งแรง ความยืดหยุ่น ความคล่องแคล่ว
 4. ทักษะทางการเคลื่อนไหว (Skilled Movement) เป็นทักษะที่ปรับจากขั้นที่งานยากขึ้น และซับซ้อน
 5. การสื่อความหมายที่ไม่อ้อมค้อม (Nondiscursive Communication) เป็นการแสดงออกเพื่อสื่อความหมาย โดยการเคลื่อนไหว หรือการเคลื่อนไหวนั้นเป็นการแสดงออกที่มีความหมายอยู่ในตัว
- เขียน ไชยศรี (2529 : 38 – 54) ได้อธิบายเกี่ยวกับพฤติกรรมการเรียนรู้ด้านการปฏิบัติว่า สามารถพิจารณาจากลักษณะต่อไปนี้ คือ

1. การเลียนแบบ (Imitation) เป็นการทำตามทีละขั้น ไปตามทีแสดงให้ดู อาจมีช่วยเหลือในขณะทำตาม โดยเน้น การทำตามแบบ ทำไปตามขั้น ทีละขั้น มีผู้ทำให้ดูหรือแสดงทีละขั้นตอนและมีการช่วยเหลือในขณะปฏิบัติ
2. การทำโดยมีการยึดแบบ (Patterning) เป็นการทำได้ด้วยตนเอง โดยการบอกแนว ให้คำชี้แจง หรือทบทวนการปฏิบัติให้ก่อน ผู้ปฏิบัติอาจทำโดยการลองผิดลองถูกตนเอง โดยเน้นทำหลังจากอธิบายวิธีการให้ฟัง ทำหลังจากทบทวนขั้นตอนให้ฟัง ทำหลังจากแสดงหรือปฏิบัติให้ดูและทำหลังจากให้ศึกษาจากคำสั่ง
3. การทำได้ด้วยความชำนาญ (Mastering) เป็นการทำได้ถูกต้องแม่นยำเหมาะสมกับเวลา โดยไม่มีการช่วยเหลือ ไม่มีการชี้แจง ไม่มีการแนะนำ ไม่มีการทำให้ดู หรือไม่มีการให้ดูรูปแบบใด ๆ เพียงแต่กำหนดหัวเรื่อง วิธีการให้ว่าให้ทำอะไร โดยเน้น ความถูกต้อง ความรวดเร็ว ความคงที่ ความประสานสัมพันธ์ ความอดทน ความแน่นอน ความถูกต้องตามสัดส่วนและความแข็งแรง
4. การทำในสถานการณ์ต่างๆ ได้ (Applying) เป็นการทำในสถานการณ์ต่าง ๆ สถานการณ์ใหม่ หรือสถานการณ์อื่น ๆ ที่นอกเหนือไปจากที่เคยทำมาแล้วได้ด้วยความต้องการในเวลาอันเหมาะสม โดยไม่มีการช่วยเหลือ ไม่มีการแนะนำขั้นตอน กระบวนการ หรือการปฏิบัติใด ๆ จากผู้อื่น โดยเน้น การเลือกทักษะที่ต้องใช้ในการแก้ปัญหา การกำหนดทักษะที่ต้องใช้ในการแก้ปัญหา ความสนใจในการใช้ทักษะนั้นในยามจำเป็นและกำหนดขั้นตอนกระบวนการในการแก้ปัญหา
5. การแก้ปัญหาโดยฉับพลัน (Improvising) เป็นการทำเพื่อแก้ปัญหาโดยฉับพลัน ซึ่งอาจจะเป็นการแก้ไขปรับปรุง เปลี่ยนแปลง ขยาย ยืดหยุ่น เสนอสดแทรกสิ่งใหม่เข้าไปกับทักษะที่มีมา หรือทำมาก่อน โดยเน้นการหาหนทางใหม่ในการใช้ทักษะเพื่อให้เหมาะสมกับสถานการณ์ การเปลี่ยนแปลงปรับปรุงทักษะที่ต้องทำในวิถีทางที่ต้องปฏิบัติ วิธีการเสริมแต่งบุคลิกบางอย่างในการที่จะปฏิบัติงานนั้นๆ และการประสานสัมพันธ์ทักษะที่ใช้ในกิจกรรมนั้น ๆ

2.3 เกณฑ์ที่ใช้ในการพิจารณาตัดสินเกี่ยวกับการวัดภาคปฏิบัติ

แบบวัดภาคปฏิบัติปกติจะใช้ทดสอบเป็นรายบุคคล การทดสอบจะใช้เวลามากทำให้สิ้นเปลือง ดังนั้นก่อนที่จะตัดสินใจใช้การทดสอบภาคปฏิบัติ ควรจะพิจารณาเกณฑ์ต่อไปนี้ (Tuckman. 1975)

1. สิ่งที่จะวัดต้องมีการปฏิบัติอย่างแท้จริง

การปฏิบัติจะเกิดขึ้นเมื่อผู้เรียนสามารถใช้มือหรือเครื่องมือทำงาน ในกรณีเช่นนี้หากใช้ข้อสอบแบบเขียนตอบวัด ถือว่าทำให้ขาดความเที่ยงตรงในการวัด แต่จะใช้การเขียนตอบในวัตถุประสงค์อื่น ๆ เช่น ถ้าต้องการวัดทักษะการคัดลายมือ ก็จำเป็นต้องอาศัยกระดาษและดินสอให้ผู้เข้าทดสอบเขียนตอบเพราะความสามารถที่มุ่งวัด คือ ความสามารถในการเขียนคัดอักษร

2. ความสามารถที่จะวัดจำเป็นต้องวัดกระบวนการปฏิบัติงาน

ในการทดสอบที่วัดกระบวนการหรือผลงาน ครูควรละเว้นจากแบบทดสอบที่ใช้กระดาษและดินสอ โดยครูต้องสังเกตการปฏิบัติงานของนักเรียนด้วยตนเอง

3. ผลงานขั้นสุดท้ายต้องอยู่ในรูปที่สัมผัสได้

ในกรณีที่วัดการตัดสินใจหรือความคิดครูสามารถใช้กระดาษและดินสอทดสอบได้ แต่ถ้าสถานการณ์ที่แตกต่างไป เช่น ถ้าผลงานที่จะวัดอยู่ในรูปของสิ่งที่มองเห็นหรือวัตถุประสงค์ของการวัด วัดทั้งความรู้ในกระบวนการและทักษะการปฏิบัติงาน ก็จำเป็นต้องใช้แบบวัดภาคปฏิบัติ

4. เมื่อต้องการที่จะประเมินการเรียนรู้ทักษะ

การวัดความรู้หรือความรู้สึกรู้สึกสามารถวัดโดยใช้แบบทดสอบกระดาษและดินสอ แต่ถ้าเป็นทักษะที่ต้องอาศัยการฝึกปฏิบัติ จนถึงแสดงทักษะที่นักเรียนจะต้องทำให้เห็น จำเป็นต้องใช้แบบวัดภาคปฏิบัติ

5. เมื่อต้องการที่จะประเมินรายบุคคลในสถานการณ์กลุ่ม

เมื่อจุดประสงค์ของการทดสอบคือ การวัดผลในเรื่องใดเรื่องหนึ่งในตัวบุคคล เช่น ทักษะความเป็นผู้นำ การทดสอบภาคปฏิบัติเป็นสิ่งจำเป็น เพราะต้องสังเกตพฤติกรรมความเป็นผู้นำในตัวผู้ถูกวัด ขณะที่อยู่ในกลุ่ม นักเรียนบางคนอาจจะเขียนอธิบายความเป็นผู้นำได้ดี แต่การปฏิบัติจริงของเขาอาจจะทำได้ไม่ดี

6. เมื่อต้องการวัดความเข้าใจในการประยุกต์

การใช้แบบวัดภาคปฏิบัติเพื่อวัดความเข้าใจที่เป็นการวัดการประยุกต์ใช้กับผลงานที่เป็นรูปธรรม แต่การปฏิบัติงานมีความชำนาญ สามารถประยุกต์ความรู้มาใช้ให้เกิดประโยชน์นั้น จำเป็นต้องอาศัยการฝึกฝนมาเป็นเวลานาน ทำให้ความสามารถในการปฏิบัติไม่ได้สะท้อนถึงความรู้ความเข้าใจ แต่เป็นความสามารถถึงขั้นที่ผู้เรียนมีทักษะการปฏิบัติที่สามารถทำได้โดยอัตโนมัติจนเป็นนิสัย เช่น ทักษะการขับรถ เป็นต้น

2.4 ชนิดของแบบทดสอบวัดภาคปฏิบัติ

การวัดด้านการปฏิบัติ นอกจากอาศัย ความรู้ ความคิดแล้ว ยังต้องอาศัยทักษะทางการปฏิบัติด้วย ซึ่งต้องพิจารณาทั้ง กระบวนการ (Process) และผลงาน (Produce) การจัดเตรียม และเวลาที่ใช้ในการปฏิบัติ จากการศึกษาวิธีการในการวัดผลด้านการปฏิบัติจากนักศึกษาลายท่านพอจะสรุปวิธีการได้ ดังนี้

มาร์แชล (Marshall. And Hales. 1971 : 139 – 141) ได้จำแนกแบบทดสอบวัดภาคปฏิบัติออกเป็น 3 ชนิด คือ

1. แบบทดสอบให้จำแนกแยะแยะ (Identification Test) มีจุดมุ่งหมายเพื่อวัดความสามารถของนักเรียนในการจำแนกวัตถุ (Object) หรือชุดของวัตถุ (Set of Objects) จำแนกหาความถูกต้องและไม่ถูกต้องในขั้นตอนการ (Procedures) และวิธีปฏิบัติ (Practices) จำแนกปัจจัยสำคัญเบื้องต้นของขั้นตอนการหรือเพื่อที่จะจำส่วนประกอบผลผลิตได้

ความแตกต่างระหว่างแบบให้จำแนกแยกแยะ และแบบทดสอบทางภาษา (Verbal Tests) บางครั้งก็ไม่ชัดเจน ถ้าให้นักเรียนระบุชื่อและชี้ส่วนต่าง ๆ ของเครื่องพิมพ์ดีด โดยทั่วไปก็เรียกว่าแบบทดสอบภาคปฏิบัติ อย่างไรก็ตามเป็นความจริงว่านักเรียนที่สามารถระบุส่วนต่างๆ ของเครื่องพิมพ์ดีดถูกต้องก็ไม่ได้เป็นหลักประกันว่านักเรียนจะมีความสามารถในการพิมพ์ การทำความสะอาดเครื่องพิมพ์ดีด หรือเปลี่ยนผ้าหมึกแบบทดสอบให้จำแนกแยกแยะไม่สามารถที่จะวัดประสิทธิภาพของผลงานในขั้นสุดท้ายของบุคคลได้

แบบทดสอบให้จำแนกแยกแยะ มักจะเกี่ยวข้องกับการท่องจำของนักเรียน ดังนั้น การจำแนกแยกแยะงาน (Identification Task) ก็ควรจะทำให้เกิดการโต้ตอบ (Reflex) ในการผสมผสานกันของทักษะและกระบวนการทางสมอง (Mental Processes)

2. แบบทดสอบจำลองสถานการณ์ (Simulated Situation Test) ในแบบทดสอบแบบจำลองสถานการณ์ กิจกรรมที่จำเป็นเกี่ยวกับงานซึ่งก็อาจจะเป็นสถานการณ์ในชีวิตที่เคยประสบมา บางครั้งผู้สอบก็มีความต้องการที่จะใช้เครื่องมือ ซึ่งถูกสร้างขึ้นมาโดยเฉพาะเพื่อที่จะฝึกหัด (Training)

โดยพื้นฐานแล้วแบบทดสอบจำลองสถานการณ์ต้องการที่จะทำการคัดเลือกกิจกรรมที่จำเป็นที่สุดเกี่ยวกับการปฏิบัติ กิจกรรมที่เคยคัดเลือกมาก็มีความมุ่งหมายเพื่อที่จะทำการจำลอง (Duplicating) หรือทำการเลียนแบบ (Simulating) ในกิจกรรมเหล่านั้น ประสิทธิภาพของแบบทดสอบนี้ขึ้นอยู่กับปริมาณของการปฏิบัติจริงที่เป็น การเลียนแบบ อย่างไรก็ตามเครื่องจักรมีราคาแพง เวลา ความสะอาด และความปลอดภัย เป็นสิ่งที่ควรพิจารณาอย่างมากในการใช้แบบทดสอบชนิดนี้

3. แบบทดสอบตัวอย่างงาน (Work Sample Test) แบบทดสอบภาคปฏิบัติจะเกี่ยวข้องกับตัวอย่างงานหลาย ๆ ชนิด เช่น ครูชวเลขให้แบบทดสอบตัวอย่างงาน เมื่อครูให้นักเรียนเขียนตัวอย่างตัวอักษรที่บอกแล้วทำการนับจำนวนของความผิดพลาดที่นักเรียนแต่ละคนทำการบันทึกไว้

อย่างไรก็ตามก็มีความยุ่งยากในการจำแนกระหว่างแบบทดสอบตัวอย่างงานและแบบทดสอบจำลองสถานการณ์ แต่ก็มีข้อสังเกตอยู่ว่าแบบทดสอบตัวอย่างงานจะใช้สถานการณ์ของงานที่เห็นจริง ซึ่งเราคาดหมายว่า จะมีความสำคัญต่อนักเรียนในการฝึกหัด และประสบการณ์ของนักเรียน ถ้าหากตัวอย่างงานถูกเลือกมาอย่างระมัดระวัง แบบทดสอบก็จะเป็นเครื่องชี้ (Indicator) ที่ดีเกี่ยวกับความสามารถของนักเรียนในการปฏิบัติกิจกรรม

กรมวิชาการ (2539 : 3-4) ได้แบ่งระดับของการวัดภาคปฏิบัติออกเป็น 4 ระดับด้วยกัน คือ

1. การทดสอบการปฏิบัติด้วยการเขียนตอบ การทดสอบแบบนี้จะแตกต่างไปจากข้อสอบเขียนตอบทั่วไป เพราะการทดสอบจะมุ่งการประยุกต์ใช้ความรู้และทักษะที่เป็นผลมาจากเรียนรู้และฝึกฝนมาก โดยคำสั่งมักจะใช้คำว่า "สร้าง" กับการสอบภาคปฏิบัติแบบนี้ การทดสอบการปฏิบัติแบบนี้จะใช้วัดคุณภาพของผลที่เกิดจากการใช้ความรู้และทักษะของนักเรียน โดยการเขียนตอบ บางครั้งก็มีความสำคัญต่อการสอบปฏิบัติมาก เช่น ในการให้นักเรียนทำกิจกรรมบางอย่างที่เสี่ยงอันตราย หรือต้องลงทุนสูง จำเป็นที่จะต้องรู้ว่าผู้ปฏิบัติมีความรู้เพียงพอหรือไม่ในการปฏิบัตินั้น ๆ จึงต้องใช้วิธีการทดสอบด้วยการเขียนตอบก่อนเพื่อที่จะช่วยหลีกเลี่ยงอันตรายหรือความผิดพลาดที่เกิดจากการปฏิบัติงานนั้น

2. การทดสอบเชิงจำแนก เป็นการทดสอบที่แยกการปฏิบัติงานออกเป็นส่วน ๆ ในระดับความเป็นจริงต่าง ๆ แล้วให้นักเรียนบอก ระบุ จำแนก หรืออธิบายในส่วนของงานนั้น เช่น ให้หาจุดที่ไฟฟ้าลัดวงจรและให้ระบุเครื่องมือ อุปกรณ์และวิธีการที่จะใช้ในการปฏิบัติงานนั้น หรือถ้าเป็นการทดสอบที่สลับซับซ้อนยิ่งขึ้นไปอีก เช่น ให้จำแนกการฟังเสียงการทำงานของเครื่องยนต์ที่ผิดปกติ แล้วหาสาเหตุว่าทำไมจึงมีเสียง

เช่นนั้น และจะแก้ไขให้เป็นปกติอย่างไร นอกจากที่กล่าวมาแล้วการทดสอบเชิงจำแนกยังใช้เป็นเครื่องมือในการเตรียมนักเรียนเพื่อการวัดในระดับการปฏิบัติจริงหรือสถานการณ์จำลอง ซึ่งจะกล่าวในลำดับต่อไป

3. การทดสอบการปฏิบัติงานในสถานการณ์จำลอง การทดสอบแบบนี้จะเน้นการให้นักเรียนได้ปฏิบัติงานในสถานการณ์ที่จำลองขึ้นมาให้มีลักษณะใกล้เคียงกับสถานการณ์จริงมากที่สุด เช่น การให้นักเรียนถอนกิ่งพืชที่ตัดออกมาจากต้นแล้ว เป็นต้น บางครั้งการจำลองสถานการณ์จะช่วยป้องกันไม่ให้นักเรียนได้รับอันตราย หรือทำให้เครื่องมือราคาแพงเสียหายระหว่างการฝึกปฏิบัติในระยะแรก ๆ การทดสอบแบบสถานการณ์จำลองนี้ บางครั้งอาจจะนำมาใช้ในการประเมินขั้นสุดท้ายของการวัดทักษะการปฏิบัติ

4. การทดสอบปฏิบัติจริง การทดสอบแบบนี้ถือว่ามีความเป็นจริงในการปฏิบัติงานอย่างสูงสุด นักเรียนจะต้องปฏิบัติงานภายใต้สภาวะการณ์จริง ซึ่งนักเรียนอาจพบปัญหาในระหว่างปฏิบัติและต้องแสดงความสามารถในการแก้ปัญหาที่พบด้วย การทดสอบแบบนี้อาจกำหนดให้นักเรียนทำงานเป็นโครงการตั้งแต่เริ่มจนสิ้นสุดโครงการก็ได้ โดยมีขั้นตอนการปฏิบัติตามสภาพจริง เช่น มีการออกแบบ การเลือกใช้วัสดุ การปฏิบัติการประเมินและปรับปรุงแก้ไข เป็นต้น

2.5 เครื่องมือวัดภาคปฏิบัติ

2.5.1 แบบมาตรวัดประเมินค่า (Rating Scale)

ความหมายของมาตรวัดประเมินค่า (Rating Scale)

ได้มีผู้ให้ความหมายของมาตรวัดประเมินค่าไว้ดังต่อไปนี้

กรมวิชาการ (2540 : 80) กล่าวว่า มาตรวัดประเมินค่า เป็นเครื่องมือที่ใช้วัดคุณลักษณะที่จัดระดับคุณภาพไว้ต่อเนื่อง และให้รายละเอียดข้อมูลค่อนข้างมาก

มัลลิกา ยุวระเดมีย์ (2527 : 12) กล่าวว่า มาตรวัดประเมินค่า คือเครื่องมือที่ใช้สำหรับประเมินค่าของสิ่งต่าง ๆ ในเชิงปริมาณ โดยได้กำหนดสเกลของลักษณะที่ต้องการประเมินมาให้ผู้ตอบเลือกตอบอย่างเป็นลำดับขั้นตอน

วิเชียร เกตุสิงห์ (2530 : 73) กล่าวว่า มาตรวัดประเมินค่า เป็นแบบทดสอบที่มีข้อความหรือข้อคำถามเกี่ยวกับเรื่องให้ผู้ตอบพิจารณา พร้อมกับมีคำตอบแสดงความเข้มของความเห็นเกี่ยวกับเรื่องนั้นเป็นระดับมาก - น้อย สูง - ต่ำ ส่วนใหญ่ใช้ตั้งแต่ 3 ระดับขึ้นไป

แซสโลว์ (Saslow, 1982 : 411) กล่าวว่า มาตรวัดประเมินค่าเป็นมาตรการจัดอันดับที่มีถ้อยคำ หรือตัวเลขแสดงไว้อย่างเป็นลำดับขั้น

กูต (Good, 1973 : 469) กล่าวว่า มาตรวัดประเมินค่า คือ สิ่งที่ยึดทำขึ้นเพื่อใช้ในการประเมินผลผลิตทัศนคติ หรือบุคลิกภาพของแต่ละบุคคล รูปแบบที่นิยมใช้กันเป็นแบบที่กำหนดตำแหน่งต่าง ๆ ไว้ให้เลือก

จากความหมายดังกล่าวข้างต้นสรุปได้ว่า มาตรวัดประเมินค่า เป็นเครื่องมือที่ใช้สำหรับประเมินค่าของสิ่งต่าง ๆ ในเชิงปริมาณ โดยมีการกำหนดรายการที่ต้องการประเมินไว้เป็นอย่างดีและมีช่วงเท่า ๆ กัน

รูปแบบของมาตรวัดประเมินค่า

มาตรวัดประเมินค่าเป็นเครื่องมือที่ใช้กันอย่างกว้างขวางทั้งในด้านการศึกษาหรือ โครงการต่าง ๆ จึงได้มีการพัฒนาปรับปรุงรูปแบบให้แตกต่างกันออกไป นักวัดผลหลายท่านจึงได้พยายามวิเคราะห์และจำแนกรูปแบบของมาตรวัดประเมินค่าตามลักษณะที่ปรากฏ แล้วสรุปเป็นรูปแบบต่าง ๆ ดังต่อไปนี้

กิลฟอร์ด (Guilford, 1954) ได้แบ่งประเภทของมาตรวัดประเมินค่าตามที่ใช้กันอยู่ออกเป็น 5 แบบ คือ

1. แบบตัวเลข (Numerical Scale) กำหนดตัวเลขพร้อมกับนิยามหรือคำบรรยายไว้ให้ผู้ประเมิน ดังตัวอย่างมาตรวัดที่ใช้ในการประมาณค่าของสีและกลิ่นดังต่อไปนี้

- 10 พอใจมากอย่างยิ่งที่สุด (Most pleasant imaginable)
- 9 พอใจมากอย่างยิ่ง (Most pleasant)
- 8 พอใจอย่างยิ่ง (Extremely pleasant)
- 7 พอใจพอสมควร (Moderately pleasant)
- 6 พอใจปานกลาง (Midly pleasant)
- 5 ไม่พอใจ (Indifferent)
- 4 ไม่พอใจปานกลาง (Midly impleasant)
- 3 ไม่พอใจพอสมควร (Moderately impleasant)
- 2 ไม่พอใจมากอย่างยิ่ง (Most impleasant)
- 1 ไม่พอใจมากอย่างยิ่งที่สุด (Most impleasant imaginable)

2. แบบกราฟิก (Graphic Scale) เป็นแบบที่นิยมใช้กันอย่างกว้างขวางมากที่สุดมีได้หลายรูปแบบ เน้นการจัดข้อความเข้ากับเส้นตรง เส้นดังกล่าวสามารถแบ่งออกเป็นช่วง ๆ หรือเป็นเส้นต่อเนื่องก็ได้ ถ้ากรณีแบ่งเป็นช่วง ๆ สามารถแบ่งได้หลายส่วน เส้นตรงทั้ง 2 แบบสามารถจัดวางตามแนวนอนหรือแนวตั้งได้ตามความเหมาะสม

3. แบบมาตรฐาน (Standard Scale) เป็นแบบที่ได้กำหนดค่าสเกล เป็นค่าประจำของชุดมาตรฐานที่จะใช้ในการเปรียบเทียบไว้ล่วงหน้า ผู้ประเมินจะต้องทำการเปรียบเทียบรายการที่ให้ประเมินกับชุดมาตรฐานมาตรวัดที่จัดอยู่ในรูปแบบนี้ได้แก่ วิธีจับคู่เหมือน (Portrait – Matching Scale) วิธีเปรียบเทียบคนต่อคน (Man – to – Man Scale)

4. การประมาณค่าโดยใช้แต้มสะสม (Rating by Cumulative Points) แตกต่างจากแบบอื่นที่วิธีการให้คะแนน วิธีการตรวจรายการ (Check – list) เป็นตัวอย่างหนึ่งของรูปแบบนี้ได้มีผู้ใช้วิธีการตรวจรายการในการประเมินคุณลักษณะของเด็ก โดยกำหนดรายการที่มีทั้งคุณลักษณะที่ดีและไม่ดีมาให้เลือก เช่น ให้ความร่วมมือ ซ้ำคิด ดูร้าย เห็นแก่ตัว เป็นต้น ผู้ประเมินแต่ละคนจะเลือกทุก ๆ คุณลักษณะที่ตรงกับเด็กคนนี้ คะแนนของเด็กแต่ละคนจะเป็นผลรวมของคะแนนทั้งหมด ซึ่งได้กำหนดให้ 1 สำหรับทุกคุณลักษณะที่ดี และให้คะแนน -1 สำหรับทุกคุณลักษณะที่ไม่ดี นอกจากนี้แบบตรวจรายการอาจจะจัดอยู่ในรูปแบบตัวเลือก (Multiple - Choice) ก็ได้

5. การประมาณค่าแบบบังคับเลือก (Forced – Choice Rating) เป็นแบบที่พัฒนาขึ้นเพื่อประเมินบุคคล ผู้ประเมินไม่ต้องบอกว่าผู้ถูกประเมินมีลักษณะใด มากน้อยเพียงใด แต่ให้เปรียบเทียบเป็นคู่ ๆ แล้วบอกว่าลักษณะใดมากกว่ากัน แต่ละข้อจะประกอบด้วยข้อความคู่ 2 คู่ ข้อความคู่แรกจะเป็นไปในทางบวก ส่วนข้อความคู่ที่เหลือจะเป็นไปในทางลบ การที่จัดรูปแบบดังกล่าวก็เพราะว่าในบางครั้งผู้ประเมินไม่ได้มี

จุดมุ่งหมายที่จะเลือกข้อความในทางบวก แต่ก็จะทำให้ชะงักที่จะเลือกข้อความในทางลบ เพราะยังมีข้อความในทางบวกให้พิจารณาอีก วิธีประเมิน ผู้ประเมินจะเลือกข้อความอันใดอันหนึ่งที่เหมาะสมหรือตรงกับลักษณะของผู้ถูกประเมิน และเลือกอีกข้อความหนึ่งที่ไม่เหมาะสมหรือไม่ตรงกับลักษณะของผู้ถูกประเมิน

ทอลล์และฮอว์กินส์ (Tull and Hawkins, 1993 : 373–389) ได้แบ่งมาตรวัดประเมินค่าออกเป็น 2 ประเภทใหญ่ ๆ ดังนี้ คือ

1. แบบไม่มีเกณฑ์เปรียบเทียบให้ (Noncomparative Rating Scale) แบบนี้ ผู้ตอบจะประเมินสิ่งเร้านั้นตามความรู้สึกนึกคิดของตนเปรียบเทียบกับเกณฑ์อะไรก็ได้ที่ผู้ตอบแต่ละคนจะเห็นว่าเหมาะสม ต่างคนต่างกำหนดเอง ไม่เป็นมาตรฐานเดียวกัน แบ่งเป็น 2 ชนิด ได้แก่

1.1 มาตรวัดประเมินค่าแบบกราฟิก (Graphic Rating Scale) ชนิดนี้พยายามทำให้การประเมินมีลักษณะต่อเนื่อง โดยให้ผู้ตอบประเมินสิ่งเร้าด้วยการทำเครื่องหมายลงบนเส้นกราฟ มี 2 ลักษณะคือ แบบเป็นเส้นกราฟตรงและมีเฉพาะตัวหนังสือบรรยายคุณลักษณะกำกับหัวท้ายเท่านั้น กับแบบเป็นเส้นกราฟตรง มีตัวหนังสือบรรยายคุณลักษณะกำกับไว้เป็นระยะ ๆ

มาตรวัดประเมินค่าด้วยเส้นกราฟนี้ เส้นตรงที่กำหนดให้ในทุกข้อจะต้องมีความยาวเท่ากัน และไม่มีการแบ่งช่วงบนเส้นไว้ให้ก่อน ผู้ตอบจะเลือกกาทลงบนเส้นกราฟ ณ จุดใดก็ได้ ตามแต่ผู้ตอบจะคาดคะเนเอาเอง

1.2 มาตรวัดประเมินค่าด้วยคำตอบจำกัด (Itemized Rating Scale) ชนิดนี้เป็นแบบกำหนดข้อความ คำ ตัวเลข หรือสัญลักษณ์ไว้เป็นคำตอบให้เลือกแน่นอน มี 3 ลักษณะ ได้แก่

1.2.1 แบบกำหนดตอบเป็นข้อความ ลักษณะนี้จะเลือกคำหรือข้อความที่มีลักษณะการเปรียบเทียบกันมาเรียงตามลำดับมากน้อยไว้ให้เลือกตอบ

1.2.2 แบบกำหนดคำตอบเป็นเส้นกราฟ ลักษณะนี้กำหนดเป็นกราฟเส้น มีคำบรรยายกำกับหัวท้าย แต่เส้นกราฟนั้นแบ่งเป็นคำตอบกำหนดไว้แน่นอน และจะมีตัวเลขกำกับไว้ด้วยก็ได้

1.2.3 แบบกำหนดคำตอบเป็นร้อยละ ลักษณะนี้กำหนดปริมาณคำตอบไว้เป็นร้อยละโดยปริมาณ โดยแบ่ง 100% เป็นช่วง ๆ เท่า ๆ กัน และมีข้อความหรือคำกำกับหัวท้าย

1. แบบมีเกณฑ์เปรียบเทียบ (Comparative Rating Scale) แบบนี้จะกำหนดเกณฑ์ไว้ให้เปรียบเทียบ ผู้ตอบทุกคนต้องใช้เกณฑ์เปรียบเทียบเดียวกัน แบ่งเป็น 4 ชนิด ได้แก่

1.1 มาตรวัดประเมินค่าแบบกราฟิก และคำตอบจำกัดในลักษณะเปรียบเทียบ (Graphic and Itemized Comparative Rating Scale) ชนิดนี้ก็คือแบบที่ไม่มีเกณฑ์เปรียบเทียบให้แน่นอน เพียงแต่เพิ่มเกณฑ์ที่ใช้เปรียบเทียบมาให้อยู่เท่านั้น

1.2 มาตรวัดประเมินค่าแบบเปรียบเทียบรายคู่ (Paired Comparisons) แบบนี้เป็นวิธีการเปรียบเทียบของ 2 สิ่งในเวลาเดียวกัน และผู้ตอบจะต้องเลือกเห็นด้วยกับข้อความใดข้อความหนึ่ง ดังนั้นคำตอบที่กำหนดให้จะต้องเป็นไปตามจริง

จากข้อความ ก กับ ข นักเรียนเห็นด้วยกับข้อความใดมากกว่า

ก. วิชาคณิตศาสตร์มีความสำคัญมากกว่าวิชาสังคมศึกษา

ข. วิชาสังคมศึกษามีความสำคัญมากกว่าวิชาคณิตศาสตร์

ในบางกรณี หากต้องการเปรียบเทียบมากกว่า 2 สิ่ง เช่นต้องการเปรียบเทียบถึง 5 วิชา ($n = 5$) ก็จะมีข้อความให้เลือกถึง 10 คู่ ดังนั้นการเปรียบเทียบรายคู่จึงมีข้อจำกัดว่าจะเปรียบเทียบแค่ 2–3 สิ่งเท่านั้น

1.3 มาตรฐานวัดประเมินค่าแบบจัดอันดับ (Rank Order Rating Scale) แบบนี้คล้ายกับการเปรียบเทียบรายคู่ เพียงแต่ให้เปรียบเทียบพร้อมกันทั้งหมด

1.4 มาตรฐานวัดประเมินค่าแบบรอยยิ้ม (Smiling Face Scale) แบบนี้จะเหมาะสมสำหรับเด็กอายุประมาณ 5 ขวบ โดยจะมีรูปใบหน้าแสดงอาการต่างๆ กำหนดให้ จากนั้นครูจะอธิบายวิธีปฏิบัติและถามคำถาม รัมเมล (Rummel, 1964 : 201 – 202) ได้กล่าวถึงมาตรฐานวัดประเมินค่าว่า ที่เป็นที่นิยมใช้มีหลายรูปแบบ พร้อมทั้งยกตัวอย่างไว้ 3 รูปแบบ ดังต่อไปนี้

1. มาตรฐานวัดประเมินค่าแบบตัวเลข (numerical Rating Scale) ให้ตัดสินลักษณะออกมาเป็นตัวเลข เช่น ให้ประเมินบุคคลตามข้อกระทงต่อไปนี้ โดยวงกลมล้อมรอบ 1 เมื่อไม่พอใจ 2 เมื่ออยู่ในระดับต่ำ 3 เมื่ออยู่ในระดับปานกลาง 4 เมื่อดีกว่าปานกลาง และ 5 เมื่ออยู่ในระดับที่ดีกว่า

2. มาตรฐานวัดประเมินค่าแบบเส้นตรง (Graphic Rating Scale) ประกอบด้วยเส้นตรงสำหรับให้ผู้ประเมินค่าทำเครื่องหมายเพื่อแสดงตำแหน่งที่เป็นจริงที่สุดเกี่ยวกับผู้ถูกประเมิน

3. มาตรฐานวัดประเมินค่าแบบบรรยาย (Descriptive Rating Scale) มักเป็นคำบรรยายแสดงระดับต่าง ๆ ของบุคลิกลักษณะ ดังตัวอย่างการประเมินพฤติกรรมโดยให้เลือกคำหรือข้อความที่บรรยายถึงพฤติกรรมของนักเรียนตามที่สังเกตได้

วอคเคิลล์ (Vockell, 1995 : 143) ได้ให้ตัวอย่างของมาตรฐานวัดประเมินค่า ไว้ดังต่อไปนี้

แบบที่ 1 ให้เขียนตัวเลขลงในช่องว่างหน้าพฤติกรรมที่กำหนดมาให้

แบบที่ 2 เป็นแบบธรรมดาเช่นเดียวกับแบบแรก แต่ให้ผู้ประเมินเลือกตัวเลขจากที่ให้ไว้

แบบที่ 4 เป็นแบบกราฟิกเช่นกัน แต่ให้ทำเครื่องหมายลงบนส่วนใดก็ได้ของเส้นตรง

วันทนี ชูคิลปี (2524 : 221) ได้แบ่งมาตรฐานวัดประเมินค่าออกเป็น 2 แบบ คือ

แบบที่ 1 ข้อคำถามที่ให้ประเมิน เป็นข้อความที่บอกอันดับมากน้อย

แบบที่ 2 ข้อคำถามที่ให้ประเมิน เป็นค่าตัวเลข

อนันต์ ศรีโสภณ (2521 : 213–215) ได้แบ่งมาตรฐานวัดประเมินค่าออกเป็น 5 รูปแบบ ดังต่อไปนี้

1. แบบตัวเลข (Numerical Rating Scale) ผู้ประเมินเพียงแต่ทำเครื่องหมายลงบนตัวเลขที่แสดงระดับหรือคุณภาพของการกระทำ หรือลักษณะนิสัย

2. แบบกราฟิก (Graphic Rating Scale) สามารถทำเครื่องหมายลงที่ใดก็ได้ตามความรู้สึกมากน้อย

3. แบบเปรียบเทียบ (Comparative Rating Scale) เป็นการเปรียบเทียบคุณภาพของสิ่งที่จะวัดกับเกณฑ์มาตรฐานที่ผู้ประเมินกำหนดขึ้น เพื่อดูความแตกต่างระหว่างตัวอย่างที่จะวัดกับเกณฑ์

4. แบบจัดอันดับ (Ranking) ให้จัดลำดับผู้ถูกประเมินจากลักษณะนิสัยสูงลงมาต่ำ เพื่อให้การจัดลำดับมีความเที่ยงมากยิ่งขึ้น ควรจัดลำดับจากสูงสุดและต่ำสุดเข้าหาที่กลาง การจัดลำดับจะช่วยให้เห็นความแตกต่างของลักษณะต่าง ๆ ได้ชัดเจนขึ้น

5. แบบเปรียบเทียบคู่ (Paired Comparison) ต้องเปรียบเทียบผู้ถูกประเมินแต่ละคนกับคนอื่นทุกคนทีละคู่ ถ้ามีลักษณะดีกว่าก็ทำเครื่องหมายไว้ เมื่อครบทุกคนแล้ว ก็นับจำนวนความถี่ที่มีลักษณะดีกว่าว่ามีจำนวนเท่าใด ด้อยกว่ามีเท่าใด

จากรูปแบบต่าง ๆ ของมาตรวัดประเมินค่าตามที่ได้กล่าวมา จะพบว่ามีลักษณะร่วมกันอยู่ 3 แบบ ได้แก่

1. แบบตัวเลข เป็นแบบที่ประกอบด้วยรายการพฤติกรรมและคำตอบที่เป็นตัวเลข โดยให้ค่าความหมายตัวเลขไว้ก่อนแล้ว ให้ผู้ประเมินเลือกตัวเลขที่ตรงกับพฤติกรรมที่สังเกตได้
2. แบบกราฟิก เป็นแบบที่คำตอบจะเป็นเส้นกราฟ มีคำหรือข้อความกำกับหัวท้ายเพื่อเป็นแนวทางในการตัดสินใจ โดยให้ผู้ประเมินทำเครื่องหมายลงบนเส้นตรง ณ ตำแหน่งใดก็ได้ เพื่อชี้ค่าที่ต้องการประเมิน
3. แบบบรรยาย เป็นแบบที่มีคำบรรยายสั้น ๆ ในแต่ละทางเลือกเพื่อประกอบการตัดสินใจให้ผู้ประเมินเลือกข้อความที่ตรงกับความรู้สึกของเขามากที่สุด

2.6 การตรวจให้คะแนนการทดสอบภาคปฏิบัติ

การเลือกรูปแบบในการให้คะแนนการสอบภาคปฏิบัติ นั้น เราควรทำก่อนที่จะออกข้อสอบ จะต้องคำนึงถึงวิธีการที่มีประสิทธิภาพ และมีความเชื่อมั่นสูงมาใช้ในการตรวจให้คะแนน (เชาวนา ชวลิตธำรง, 2534 : 45-50) ได้กล่าวถึงการตรวจให้คะแนนของการทดสอบ ดังนี้

1. การตรวจให้คะแนนการทดสอบการจำแนก เป็นการตรวจให้คะแนนเช่นเดียวกับแบบถูก-ผิด จับคู่ เลือกตอบ เติมคำ หรือบรรยายตามชนิดของข้อสอบที่ใช้ในการทดสอบ สำหรับเด็กเล็ก ๆ หรือผู้สอบที่เขียนหนังสือไม่ได้ก็จำเป็นต้องใช้แบบปากเปล่า ในกรณีเช่นนี้ ขณะดำเนินการสอบ ครูหรือผู้ดำเนินการสอบ จะต้องบันทึกคำตอบนักเรียนแต่ละคำตอบของนักเรียนแต่ละคนลงไปในกระดาษคำตอบที่เตรียมไว้

การให้คะแนน ถ้านักเรียนตอบโดยบันทึกคำตอบในแต่ละข้อด้วยข้อความหนึ่ง หรือประโยค ก็ให้คะแนนโดยใช้วิธีเดียวกับการให้คะแนนแบบบรรยายสั้น ๆ ซึ่งอาจจะให้คะแนนเป็นถูก 1 ผิด 0 หรือถูกต้อง สมบูรณ์ 2 ถูกไม่สมบูรณ์ 1 ผิด 0 เป็นต้น การกำหนดคะแนนเต็มในแต่ละข้ออาจไม่เท่ากันก็ได้ขึ้นอยู่กับน้ำหนักความสำคัญและความเหมาะสม

2. การตรวจให้คะแนนการทดสอบสถานการณ์จำลอง และตัวอย่างงาน วิธีการให้คะแนน “กระบวนการทำงาน” หรือ “ผลงาน” เป็นสิ่งที่มีความสำคัญมากต่อการแปลผลการทดสอบการปฏิบัติให้มีความหมายที่ถูกต้อง จึงจำเป็นที่จะต้องเลือกใช้ /สร้างเครื่องมือ ด้วยความระมัดระวัง การสร้างเครื่องมือจะต้องพิจารณาถึงความเชื่อมั่นในการวัด ความสะดวกสำหรับผู้ผู้ใช้ และความเหมาะสมของชิ้นงาน เครื่องมือที่ใช้กันมากในการให้คะแนน ได้แก่ แบบตรวจสอบ (Checklists) มาตรฐานประมาณค่า (Rating Scales) และแบบบันทึกย่อ (Anecdotal Forms)

การให้คะแนนโดยใช้แบบตรวจสอบ และมาตรฐานประมาณค่า อยู่ในรูปของมาตรวัด (Scales) แบบต่าง ๆ ดังนี้

1. มาตรากการวัดเชิงปริมาณ (Quantitative Scales) เป็นการให้คะแนนในรูปความถี่ของพฤติกรรมที่ปรากฏออกมา ครูผู้ให้คะแนนจะต้องสังเกตพฤติกรรมในกระบวนการทำงาน หรือผลงานของนักเรียน และบันทึกความถี่ในแบบฟอร์มที่จัดเตรียมไว้สำหรับงานนั้น ๆ มาตรากการวัดส่วนใหญ่ จะมีระหว่าง 2 ถึง 5 ระดับ
2. มาตรากการวัดเชิงคุณภาพ (Qualitative Scales) เป็นการพิจารณาพฤติกรรมของนักเรียนในรูประดับคุณภาพของงานว่า ดีมากน้อยเพียงใด มักนิยมใช้ 3 หรือ 5 ระดับ

3. มาตรการวัดสถานภาพ (Status Scales) เป็นการพิจารณาในรูป “ความดี” ของพฤติกรรมของนักเรียน เปรียบเทียบกับกลุ่มอ้างอิง แล้วตัดสินว่า “สูงกว่า” หรือ “ต่ำกว่า” ระดับเฉลี่ยในกลุ่มอ้างอิง โดยทั่วไปนิยมใช้ 2 หรือ 5 ระดับ

4. มาตรการวัดเชิงพรรณนา (Descriptive Scales) เป็นการให้คะแนนตามระดับการปฏิบัติงาน หรือลักษณะของงานตามที่บรรยายไว้

5. มาตรการวัดผลงาน (Product Scales) เป็นการให้คะแนนผลงานของนักเรียนโดยนำไปเทียบกับผลงานที่เป็นเกณฑ์ในแต่ละระดับ

ในการกำหนดมาตรฐานการประเมินผลการปฏิบัติงานนี้ จะต้องมิเกณฑ์หรือมาตรฐานที่จะใช้เป็นหลักการพิจารณาตัดสิน การกำหนดมาตรฐานไม่ได้ดูเพียงแต่ละลักษณะเนื้อแท้ของกิจกรรม จะต้องดูที่อายุของนักเรียนและจุดประสงค์ของวิชาด้วย ในการนิยามมาตรฐานการปฏิบัติงานมีสิ่งตีพิมพ์จำนวนมากที่ช่วยในการกำหนดมาตรฐานได้อย่างเที่ยงตรง เช่น หนังสือแบบเรียน ตำราต่าง ๆ คู่มือครู และเอกสารที่จัดพิมพ์โดยสมาคมอาชีพนั้น ๆ สิ่งเหล่านี้สามารถนำมาประกอบในการกำหนดเกณฑ์ให้ได้มาตรฐานแน่นอน

2.7 เกณฑ์การให้คะแนนผลงาน (Scoring Rubric)

กรมวิชาการ (2539 : 54-59) ได้ให้ความหมายไว้ว่า แนวทางในการให้คะแนน (Scoring Rubric) ซึ่งจะต้องกำหนดเป็นมาตรวัด (Scale) และรายการคุณลักษณะที่บรรยายถึงความสามารถในการแสดงออกของแต่ละจุดในมาตรวัดอย่างชัดเจน จึงมีส่วนสำคัญในการส่งเสริมการเรียนรู้ โดยทำให้เป้าหมายการแสดงออกของนักเรียนมีความชัดเจนขึ้นนำไปสู่การบรรลุจุดประสงค์หรือสมรรถภาพที่สำคัญของมาตรฐานการศึกษาได้

เสาวนีย์ เกียรติ (2540 : 159) กล่าวว่า เป็นเครื่องมือในการให้คะแนนที่ประกอบด้วยประเด็นต่าง ๆ (Criteria) ที่จะใช้พิจารณางานหนึ่ง ๆ และคำอธิบายระดับคุณภาพของแต่ละประเด็นประเมิน ซึ่งอาจเรียกลำดับตั้งแต่ดีเลิศไปจนถึงต้องปรับปรุง หรือให้เป็นระดับตัวเลขตั้งแต่มากที่สุด (เช่น 4) ไปจนถึงน้อยที่สุด (เช่น 0) ประเด็นประเมินอาจกำหนดเพิ่มเติมได้หลายข้อ คำอธิบายระดับคุณภาพควรอธิบายให้ชัดเจนที่สุดระดับที่สุด เป็นคำอธิบายที่สามารถบอกได้ว่า ทำไมจึงต้องดีเลิศ ดี ต้องปรับปรุง

สเตนมาร์ค (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 รายละเอียดจะเปลี่ยนแปลงสำหรับระดับคะแนนที่แตกต่าง แนวคิดพื้นฐานสามารถประยุกต์ใช้ได้ในทุกระดับ

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

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

จากผลการศึกษาดังกล่าว ผู้วิจัยขอให้ความหมายของ กฎเกณฑ์การให้คะแนน (Scoring Rubric) หมายถึง แนวทางที่กำหนดขึ้นเพื่อชี้บอกระดับของพฤติกรรม คุณภาพของกระบวนการทำงานและผลลัพธ์ หรือสิ่งที่ต้องการ ซึ่งได้มาจากการพิจารณาผลการประเมินของผู้ประเมินให้คะแนน ทำให้เกิดความเข้าใจตรงกัน

2.8 ประเภทของกฎเกณฑ์การให้คะแนน

นิตโก้ (Nitko, 1983 : 266) ได้จัดประเภทของกฎเกณฑ์การให้คะแนน (Scoring Rubric) ออกเป็น 3 ประเภท คือ

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

2. Analytic Rubric ประเภทนี้ใช้เวลาในการประเมินมากกว่าในแบบ Holistic Rubrics ผู้ประเมินจะต้องพิจารณาและประเมินทีละด้าน รายละเอียดเหล่านี้เป็นสิ่งที่มีความสำคัญ ต่อการนำไปใช้หรือวินิจฉัยเด็กให้เข้าใจถึงสิ่งที่คาดหวังในการปฏิบัติแต่ละด้าน มีประโยชน์ช่วยในการเรียนรู้ของนักเรียนแม้จะใช้เวลาามากก็ตาม

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

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

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

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

- 2.1 ความเข้าใจในความคิดรวบยอด ข้อเท็จจริง เป็นการแสดงให้เห็นว่านักเรียนเข้าใจในความคิดรวบยอด หลักการในการแก้ปัญหาที่ตามอย่างกระฉับกระช่าย
- 2.2 การสื่อความหมาย สื่อสาร คือ ความสามารถในการอธิบาย นำเสนอ การบรรยายเหตุผล แนวคิด ให้ผู้อื่นเข้าใจได้ดี มีความคิดสร้างสรรค์
- 2.3 การใช้กระบวนการและยุทธวิธี สามารถเลือกยุทธวิธี กระบวนการนำไปสู่ความสำเร็จอย่างมีประสิทธิภาพ
- 2.4 ความสำเร็จของงาน ความถูกต้องแม่นยำในผลสำเร็จของงานหรืออธิบายที่มาและการตรวจสอบผลงาน

บุญเชิด ภิญญอนันตพงษ์ (2544 : 90) กล่าวว่า การพิจารณากฎเกณฑ์การให้คะแนน 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

2.9 การสร้างกฎเกณฑ์การให้คะแนน

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

ขั้นที่ 1 วิเคราะห์ผลการเรียนรู้ของนักเรียน

วิเคราะห์ผลการเรียนรู้ของนักเรียนในแต่ละจุดประสงค์หรือ แต่ละหัวเรื่อง (Theme) เมื่อนักเรียนได้ปฏิบัติกิจกรรมการเรียนรู้แล้วจะเกิดผลการเรียนรู้อะไรบ้าง ผลการเรียนรู้ของนักเรียนไม่จำเป็นจะต้องมีครบทุกประเภท ทุกครั้งที่นักเรียนได้ปฏิบัติกิจกรรมการเรียนรู้ บางครั้งในการวางแผนการสอน ครูคาดหวังในผลการเรียนรู้ของนักเรียนเพียงกระบวนการอย่างเดียว บางครั้งครูคาดหวังทั้งกระบวนการและผลงาน

ตัวอย่าง

การพูดอภิปรายเกี่ยวกับเรื่องราวต่าง ๆ ในชีวิตประจำวันผลการเรียนรู้ของนักเรียนที่ครูคาดหวัง

1. การวางแผนการจัดอภิปราย
2. การกำหนดเค้าโครงและรูปแบบการอภิปราย
3. การพูดอภิปรายที่ดี

ขั้นที่ 2 กำหนดประเด็นที่จะต้องประเมิน

การกำหนดประเด็นที่ต้องการประเมิน อาจให้นักเรียนร่วมกันเสนอความคิดในการกำหนดประเด็น ประเมินผลการเรียนรู้แต่ละอย่างของเขาเอง เช่น ครูกำหนดว่า ถ้าเราจะดูคนดีจะดูอย่างไรบ้าง ให้นักเรียนช่วยกันเสนอประเด็นที่จะดูคนดี ซึ่งมีประเด็นมากมายเป็นของนักเรียนเอง แต่อย่างไรก็ตามเนื่องจากในการวางแผนการสอน ครูผู้สอนได้มีการคาดหวังความสำเร็จจากการเรียนรู้ของนักเรียนเอาไว้ล่วงหน้าแล้ว จึงควรกำหนดรายการประเมินที่สำคัญ ๆ ของความสำเร็จจากการเรียนแต่ละด้านเอาไว้ล่วงหน้าด้วย เพื่อจะได้แจ้งให้นักเรียนทราบและเพื่อการตรวจสอบผลงานของนักเรียนเอง

ตัวอย่าง ประเมินการอภิปราย

รายการประเมิน

1. การแสดงความคิดเห็น
2. เนื้อหาสาระ
3. การนำเสนอประเด็นอภิปราย
4. การใช้ถ้อยคำ
5. มารยาทในการพูด
6. การรักษาเวลา

ขั้นที่ 3 คัดเลือกประเด็นประเมินผลที่สำคัญ

การกำหนดประเด็นที่จะต้องประเมิน เราจะพบว่า ในการผลการเรียนรู้หนึ่งอย่าง จะมีประเด็นที่ต้องการประเมินมาก โดยเฉพาะถ้าเปิดโอกาสให้นักเรียนมีส่วนร่วมในการเสนอประเด็นประเมินด้วยแล้ว จะมีมุมมองในการประเมินที่หลากหลาย หรือรวมประเด็นที่สำคัญ ๆ

ขั้นที่ 4 เลือกรูปแบบในการสร้างเกณฑ์

การประเมินการสร้างเกณฑ์การประเมินผลการเรียนรู้ของนักเรียน เพื่อพิจารณาตัดสินให้คะแนน (Scoring Rubric) มีรูปแบบในการสร้าง 2 รูปแบบ

รูปแบบที่ 1 การให้คะแนนแบบรวมองค์ประกอบ (Holistic Scoring Rubric)

การสร้างเกณฑ์การประเมินแบบองค์ประกอบ หมายถึง การให้คะแนนผลการเรียนรู้โดยรวมทุกประเด็นที่กำหนดเพื่อการประเมิน แล้วเขียนอธิบายคุณภาพของผลการเรียนรู้แต่ละระดับ

ตัวอย่าง รูปแบบการให้คะแนนเป็นภาพรวม

ประเมิน ความสามารถในการอ่านจับใจความสำคัญ

ประเด็นประเมิน

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

ตาราง 3 เกณฑ์การให้คะแนนแบบรวมองค์ประกอบ

ระดับคุณภาพ	คำอธิบาย
ระดับ 3	หมายถึง ตอบคำถามจากเรื่องที่อ่านถูกต้อง 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 จุดด้อยของแต่ละประเมินประเด็นได้ชัดเจน การสร้างเกณฑ์การประเมินในรูปแบบนี้จะต้องเขียนคำอธิบายคุณภาพของงานในแต่ละองค์ประกอบ และแต่ละระดับขององค์ประกอบให้ชัดเจน

ตัวอย่าง รูปแบบการให้คะแนนแบบแยกองค์ประกอบ

ประเมิน ความสามารถในการอ่านจับใจความสำคัญ

ประเด็นประเมิน

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

ตาราง 4 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	- บอกเนื้อหาสาระถูกต้อง แต่รวน	- บอกเนื้อหาสาระได้บ้าง
3. การบอกข้อคิดจากเรื่องที่อ่าน	- บอกข้อคิดได้ตรงประเด็นสมบูรณ์	- บอกข้อคิดได้ตรงประเด็นแต่ไม่ต่อเนื่อง	- บอกข้อคิดได้บ้างแต่รวน
4. การเสนอความคิดเห็นจากเรื่องที่อ่าน	- เสนอความคิดเห็นด้วยเหตุผลและประโยชน์	- 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 ในกรณีกำหนดค่าระดับต่ำสุดที่เลข 0 นั้นหมายถึง นักเรียนไม่มีผลการเรียนรู้ หรือไม่มี ผลงานเข้ารับการประเมินหรือไม่ได้ดำเนินการตามรายการประเมินนั้น และการกำหนดค่าระดับจาก 0-10 แสดงว่าผู้กำหนดค่าระดับคุณภาพ พิจารณาแล้วเห็นว่าคุณภาพของงานหรือสิ่งที่จะประเมินควรมีคะแนนสูงสุดคือ 10 เช่น การกำหนดค่าระดับในการประเมินการแข่งขันทักษะทางวิชาชีพ ระดับเขตการศึกษา ผู้เข้าแข่งขันเป็นผู้ที่ได้รับการคัดเลือกจากจังหวัดมาแล้ว ถือว่าเคยผลิตผลงานที่มีคุณภาพมาก่อน

ขั้นที่ 6 บรรยายคุณภาพการประเมินแต่ละระดับ

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

ตาราง 5 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 การออกเสียงไม่ถูกต้องชัดเจน แต่ไม่รักษาเวลาและมารยาทในการพูด
1 ใช้ไม่ได้	หมายถึง มีการนำเสนอข้อเท็จจริงและข้อคิดเห็น แต่ขาดการยกตัวอย่างเพื่อสนับสนุนประเด็นอภิปราย การออกเสียงไม่ถูกต้องชัดเจน ขาดการรักษาเวลาและมารยาทในการพูด
0	หมายถึง ไม่มีการอภิปราย

ขั้นที่ 7 กำหนดคะแนนการตัดสินระดับคุณภาพ

การประเมินผลงานหรือชิ้นงานเรามักจะได้ยินคำพูดเชิงการบ่งบอกถึงการจัดระดับคุณภาพ (เกรด) ของผลงาน เช่นพูดว่า ผลงานชิ้นนั้นดีมาก ชิ้นนั้นดี ชิ้นนั้นพอใช้แต่ไม่เคยบอกว่า คะแนนระหว่างเท่าไร ถึงเท่าไรที่แสดงว่าผลงานดีมาก การกำหนดคะแนนการตัดสินระดับคุณภาพ เป็นการกำหนดช่วงคะแนนจากการประเมินผลการเรียนรู้

ตาราง 6 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. สร้างเกณฑ์การประเมินในลักษณะของมาตรฐานจัดอันดับคุณภาพ ซึ่งประกอบด้วยเกณฑ์การประเมินผลงานในแต่ละทักษะย่อย
6. นำเกณฑ์การประเมินไปให้ผู้เชี่ยวชาญทางการวัดผลพิจารณาความเที่ยงตรงเชิงเนื้อหาตามวิธีของโรวินสกีและเฮมเบิลตัน
7. เลือกตัวอย่างผลงานของนักเรียนที่ไม่ได้เลือกเป็นผลงานดีเด่นมาทักษะละ 1 ผลงาน นำไปให้ผู้ตรวจให้คะแนนจำนวน 2 คน ทดลองใช้เกณฑ์การประเมินผลงานประเมินให้คะแนนผลงานจนครบทุกทักษะ
8. หาคความสอดคล้องของการให้คะแนนโดยใช้ค่าสหสัมพันธ์แบบเพียร์สัน และทดสอบความมีนัยสำคัญทางสถิติด้วย t-test

เสาวนีย์ เกียรติ (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 ให้นักเรียนลองใช้รู้จักกับงานที่ให้ดูเป็นตัวอย่างในขั้นที่ 1
 5. หัดใช้ในการประเมินตนเองและให้เพื่อนประเมินให้นักเรียนทำ ในระหว่างที่นักเรียนกำลังทำงาน ให้หยุดเป็นระยะ ๆ เพื่อประเมินตนเองและให้เพื่อนประเมิน
 6. ทบทวนให้โอกาส นักเรียนทบทวนและปรับปรุงงานของเขาโดยดูจากผลสะท้อนที่ได้จากขั้นตอนที่ 5
 7. ครูประเมิน ครูใช้รูปที่นักเรียนใช้นั้น เพื่อประเมินผลงานของนักเรียน
- บุญเชิด ภิญโญนนท์พงษ์ (2544 : 94) กล่าวว่า การสร้างกฎเกณฑ์การให้คะแนนการประเมินผลภาคปฏิบัติต้องมีส่วนประกอบในการดำเนินการที่สำคัญ 3 ขั้นตอน คือ

1. กำหนดการปฏิบัติให้ชัดเจน
 - 1.1 ระบุชนิดของการปฏิบัติว่า เป็นทักษะวิธีการทำหรือผลงานสำเร็จที่สร้างขึ้น
 - 1.2 กำหนดจุดสนใจของการประเมินว่า สนใจเป็นรายบุคคลหรือรายกลุ่ม
 - 1.3 กำหนดเกณฑ์การปฏิบัติที่สะท้อนคุณลักษณะที่สำคัญของการปฏิบัติที่ประสบผลสำเร็จ
2. ออกแบบกิจกรรมสำหรับใช้ประเมิน ดังนี้
 - 2.1 ระบุลักษณะของงานหรือกิจกรรมว่า เป็นผลงานหรือการปฏิบัติตามที่มอบหมายให้ทำหรือเป็นการแสดงออกที่เกิดขึ้นตามธรรมชาติ
 - 2.2 กำหนดรายการของงานหรือกิจกรรม ได้แก่ นิยามจุดมุ่งหมายของการปฏิบัติ เงื่อนไขสถานการณ์ของการปฏิบัติ และเกณฑ์มาตรฐาน
 - 2.3 กำหนดจำนวนชิ้นงานหรือกิจกรรมให้ทำว่าจะใช้เพียงงานเดียวหรือหลายงานจึงจะเพียงพอในสภาพการณ์นั้น
3. ให้คะแนนและจัดบันทึกผล ดังนี้
 - 3.1 ระบุรูปแบบการให้คะแนนว่าใช้เกณฑ์ในการให้คะแนนแบบใด เช่น แบบทั่วไปหรือแบบเฉพาะเนื้อหา และแบบเกณฑ์รวมหรือแบบเกณฑ์ย่อย
 - 3.2 กำหนดวิธีการจัดบันทึกว่าเป็นแบบตรวจสอบรายการ มาตรฐานประเมินค่าการปฏิบัติ การจัดบันทึกพฤติกรรมตามช่วงเวลา หรือการจัดจำรรวมค่า
 - 3.3 ระบุผู้ประเมินว่าจะให้ใครเป็นผู้สังเกตและประเมิน ผลการปฏิบัติจะเป็นประโยชน์แก่ครูผู้สอน ครูที่เคยสอน ตัวผู้เรียนเอง หรือเพื่อน ๆ ในกลุ่ม

การศึกษาค้นคว้าในครั้งนี้ ผู้วิจัยได้ใช้วิธีผสมผสานจากหลายแนวทางที่กล่าวมาข้างต้น มาสร้างเป็นกฎเกณฑ์การให้คะแนนในการประเมินผลงานศิลปะ

2.10 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. ไม่ควรย้อนไปดูคะแนนของเด็กจากข้อที่ตรวจแล้วในการตรวจข้อต่อไป ทั้งนี้ เพื่อป้องกันไม่ให้คะแนนในข้ออื่น ๆ มีผลกระทบต่อการให้คะแนนในข้ออื่น ๆ
4. ไม่ควรให้คะแนนโดยยึดความถูกต้องทางภาษาเป็นหลัก ถ้าหากมิได้มุ่งวัดความถูกต้องในการเขียนและการใช้ภาษา ความถูกต้องสละสลวยในการใช้ถ้อยคำสำนวนในการตอบ ไม่ควรมีอิทธิพลต่อการให้

คะแนนมากหรือน้อย ผู้ตรวจควรพิจารณาเฉพาะเป้าหมายการตอบในแง่ความสมบูรณ์ครบถ้วนของเนื้อหาหรือความสมเหตุสมผลของความคิดเป็นหลักในการให้คะแนน

5. ถ้าเป็นไปได้ ควรให้คนอื่นตรวจสอบผลการตรวจให้คะแนน ทั้งนี้เพื่อให้คนอื่นได้ประเมินความเหมาะสมในการให้คะแนนของเรา ตามหลักการที่ถูกต้องนั้น คะแนนที่เด็กได้ควรเป็นคะแนนเฉลี่ยที่เกิดจากผู้ตรวจให้คะแนนหลาย ๆ คน ซึ่งคงเป็นเรื่องยากในเชิงปฏิบัติ แต่อย่างไรก็ตามถ้าไม่อาจปฏิบัติได้อย่างน้อยผู้ตรวจควรจะได้ทบทวนความเหมาะสมในการตรวจให้คะแนนของตนอีกครั้งหนึ่งก่อนที่จะนำคะแนนเหล่านั้นไปใช้

6. ควรเขียนวิจารณ์ ข้อท้วงติง (Comments) เพื่อประโยชน์ในการเรียนรู้ของผู้เรียน หรืออย่างน้อยได้เป็นหลักฐานว่า ทำไมจึงให้คะแนนเท่านี้ สำหรับคำตอบของเด็กคนนี้

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

ส. วาสนา ประवालพฤกษ์ (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 โดยจะหักจากระดับคะแนนสูงสุดลงมาที่ระดับ
4. กำหนดระดับการยอมรับและคำอธิบาย
5. ใช้หลักการจัดกลุ่มแบบอิงกลุ่ม

สเทนมาร์ค (Stenmark, 1991 : 24) ได้เสนอวิธีการกำหนดเกณฑ์การให้คะแนนเพื่อใช้ในการตรวจให้คะแนนแบบประเมินผลรวม (Holistic) โดยการแบ่งข้อสอบออกเป็นกองต่าง ๆ ดังนี้ ข้อสอบที่มีคะแนนระดับสูงสุด ระดับกลาง และระดับต่ำ และสามารถเพิ่มกองในการแบ่งระดับคะแนนได้ ทุกระดับคะแนนจะอธิบายถึงการประเมินคำตอบในการแก้ปัญหาหนึ่ง ๆ โดยเฉพาะ ไม่นำมาใช้ในการจัดหมวดหมู่ของนักเรียน ระดับคะแนนในแต่ละข้อไม่จำเป็นต้องเหมือนกันกับการสร้างเกณฑ์รูปคเฉพาะ (Specific Rubric) และได้เสนอแนวทางการให้คะแนนแบบรวมองค์ประกอบ (Holistic) โดยแบ่งวิธีการให้คะแนนได้หลายวิธี ดังนี้

วิธีที่ 1 แบ่งงานตามคุณภาพเป็น 3 กอง คือ

กองที่ 1 ได้แก่งานที่มีคุณภาพเป็นพิเศษ เขียนอธิบายลักษณะของงานที่มีคุณภาพเป็นพิเศษ

กองที่ 2 ได้แก่งานที่ยอมรับได้ เขียนอธิบายลักษณะของงานที่ยอมรับได้

กองที่ 3 ได้แก่งานที่ยอมรับได้น้อยหรือยอมรับไม่ได้

จากนั้นก็ทำงานแต่ละกองมาให้คะแนนเป็น 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 โดยพิจารณาจากความบกพร่องของคำตอบว่ามีมากน้อยเพียงใด โดยหักจากระดับคะแนนสูงสุดมาที่ระดับ ดังนี้

คะแนน 4 หมายถึง คำตอบถูกแสดงเหตุผลถูกต้อง แนวคิดชัดเจน

คะแนน 3 หมายถึง คำตอบถูก เหตุผลถูกต้อง มีข้อผิดพลาดเล็กน้อย

คะแนน 2 หมายถึง เหตุผลหรือการคำนวณผิดพลาด แต่มีแนวทางที่จะนำไปสู่คำตอบ

คะแนน 1 หมายถึง แสดงวิธีคิดเล็กน้อยแต่ไม่ได้คำตอบ

คะแนน 0 หมายถึง ไม่ตอบหรือตอบไม่ถูก

วิธีที่ 3 อธิบายคุณภาพตามระดับจากสูงสุดถึงต่ำสุด เช่น เกณฑ์การให้คะแนนความสามารถในการทำความเข้าใจเนื้อหาสาระ เขียนได้ 4 ระดับดังนี้

คะแนน 4 หมายถึง การสาธิตหรือแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์ครบถ้วนถูกต้องแม่นยำในหลักการความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนด รวมทั้งเสนอแนวคิดใหม่ que แสดงถึงความเข้าใจอย่างลึกซึ้งถึงกฎเกณฑ์หรือลักษณะของข้อมูล

คะแนน 3 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจที่สมบูรณ์ ครบถ้วนถูกต้องในหลักการความคิดรวบยอด ข้อเท็จจริงของงานหรือสถานการณ์ที่กำหนด

คะแนน 2 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจไม่สมบูรณ์ ครบถ้วนถูกต้องในหลักการความคิดรวบยอด ข้อเท็จจริงของงาน หรือสถานการณ์ที่กำหนดในบางส่วน

คะแนน 1 หมายถึง การแสดงออกให้เห็นถึงการเข้าใจหลักการความคิดรวบยอด ข้อเท็จจริงหรือสถานการณ์ที่กำหนดน้อยมาก และเข้าใจไม่ถูกต้องบางส่วน

คะแนน 0 หมายถึง ไม่แสดงความคิดเห็นใด ๆ

คอร์ทนี่ (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

งานที่แสดงว่าไม่ได้ความพยายามเลยให้ 0 คะแนนและเสนอแนะประเด็นการกำหนดเกณฑ์การให้คะแนนแบบรูปทั่ว ๆ ไป (General Rubric) มีดังนี้

ระดับสูงสุด

- ประกอบด้วยคำตอบที่สมบูรณ์ ชัดเจน ปะติดปะต่อ ไม่คลุมเครือและอธิบายได้ดี
- รวมถึงแผนภาพที่ไม่ชัดเจนและซับซ้อน

- สื่อสารอย่างมีประสิทธิภาพเพื่อแสดงต่อผู้อ่าน
- แสดงความเข้าใจแนวคิดและกระบวนการของคำตอบที่เกี่ยวข้อง
- 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 คีลาเดช (2540 : 147) กล่าวว่า การตัดสินให้คะแนนนักเรียนเป็นกิจกรรมที่ประเมินความสามารถของนักเรียนจากหลักฐานผลงานมาตีค่าเป็นคะแนน การที่จะทำให้ผลการตัดสินน่าเชื่อถือนั้น มีปัจจัยสำคัญ 2 ประการ คือ

1. กฎเกณฑ์การให้คะแนน ต้องได้รับการสร้างและพัฒนาจนมีคุณภาพ โดยเฉพาะความเที่ยงตรงตามเนื้อหา นั่นคือ ต้องสร้างให้สอดคล้องกับจุดมุ่งหมายของงาน ในทางปฏิบัติควรช่วยสร้างและนำไปทดลองใช้ก่อนเพื่อหาข้อบกพร่องและปรับปรุง
2. ผู้ประเมิน เพื่อให้เกิดความน่าเชื่อถือและยุติธรรมต่อนักเรียน ควรใช้ผู้ประเมินมากกว่า 1 คน และมีการฝึกใช้กฎเกณฑ์การให้คะแนนก่อน เพื่อความชัดเจนและประเมินได้ตรงประเด็น ถ้ามีผู้สอนร่วมกันหลายคน ควรให้ทุกคนประเมินผลงานนักเรียนแต่ละคน แล้วนำเอาคะแนนของผู้ประเมินทุกคนมาเฉลี่ย ในกรณีที่มีความจำเป็นต้องประเมินเพียงคนเดียว ควรประเมินซ้ำอีกครั้งหนึ่ง โดยทั้งช่วงเวลาการประเมิน แต่ละครั้งห่างกันพอสมควร นำคะแนนทั้งสองครั้งมาเฉลี่ย

จากการศึกษาของผู้วิจัยพบว่า ถ้าวางแผนการดำเนินงานชัดเจน มีการกำหนดให้ทำกิจกรรมครบถ้วน ผลสำเร็จของงานสอดคล้องกับกฎเกณฑ์การให้คะแนนแล้ว การตัดสินให้คะแนนของผู้ประเมินจะค่อนข้างสอดคล้องกันมาก ค่าความเชื่อมั่นของการประเมินค่อนข้างสูง และเป็นไปตามสภาพที่แท้จริงของ

ผลงานศิลปะ ผู้วิจัยจึงได้กำหนดรูปแบบของมาตราประเมินให้คะแนนเป็น 2 รูปแบบ แล้วศึกษาคำความเชื่อมั่นของการประเมิน ให้คะแนนแต่ละรูปแบบโดยใช้สูตรของ P_0 และ KAPPA และ RAI

2.11 ลักษณะของศิลปะ

ศิลปะเป็นสื่อความเข้าใจระหว่างมนุษย์ ช่วยผ่อนคลายความเครียดทางอารมณ์ และพัฒนาความคิดสร้างสรรค์ให้กับผู้ปฏิบัติ ผลงานทางศิลปะนอกจากจะมีความสวยงามแล้วยังสามารถนำไปใช้ประโยชน์ได้อีกด้วย ซึ่งมีรูปแบบต่างดังนี้

1. การเขียนภาพระบายสี เป็นกิจกรรมทางศิลปะที่สนุกสนานเพลิดเพลิน เกิดจากการใช้วัสดุวาดหรือขีดเขียนเป็นรูปต่าง แล้วใช้สีระบายให้สวยงามตามจินตนาการและความคิดสร้างสรรค์ ประกอบด้วย การเขียนภาพศิลปะไทย การเขียนภาพศิลปะท้องถิ่น การเขียนภาพศิลปะสากล
2. การทดลองใช้สีด้วยเทคนิคต่าง ๆ ประกอบด้วย การใช้สีน้ำร่วมกับสีเทียน การสไลด์สี การเอียงสี การโรยสี การหยดสี เทสี พรหมสี ป้ายสี
3. การปั้นและการแกะสลัก ประกอบด้วย การปั้นภาพจากเหตุการณ์ในชุมชน การปั้นของเล่น การปั้นเพื่อประดับตกแต่ง การปั้นงานศิลปะไทย การปั้นภาพจากเหตุการณ์ การแกะสลักวัสดุเนื้ออ่อน
4. การพิมพ์ภาพ ประกอบด้วย การพิมพ์ภาพด้วยส่วนต่าง ๆ ของมือ การพิมพ์ภาพจากแม่พิมพ์ดินเหนียว การพิมพ์ภาพจากกระดาษพับเป็นลวดลายหรือรูปทรง การพิมพ์ภาพด้วยเศษวัสดุต่าง ๆ การพิมพ์ภาพโดยใช้วัสดุเนื้ออ่อนมาแกะเป็นแม่พิมพ์ (วินัย พัฒนรัฐ. 2540 : 110 -138)

2.11.1 การประเมินผลศิลปะภาคปฏิบัติ

ลักษณะการประเมินศิลปะภาคปฏิบัติ ไอส์เนอร์ (Eisner. 1972 : 208 – 209) ได้กล่าวถึงลักษณะในการประเมินผลโดยการตัดสินเปรียบเทียบกับ 3 ลักษณะ คือ

1. การเปรียบเทียบกับตนเอง เพื่อให้รู้ถึงระดับความสามารถ อารมณ์ และความเข้าใจของผู้เรียน ว่ามีการเจริญงอกงามเพิ่มขึ้นมากน้อยแค่ไหน เช่น การเปรียบเทียบผลงานศิลปะที่ทำเมื่อตอนต้นภาคเรียนกับผลงานศิลปะที่ทำเมื่อปลายภาคเรียน
2. เปรียบเทียบกับเพื่อนร่วมชั้น เพื่อให้รู้ถึงระดับความสามารถและการปฏิบัติงานของกลุ่ม การเปรียบเทียบลักษณะนี้จำเป็นต้องคำนึงถึงวุฒิภาวะของผู้เรียนด้วย เช่นการเปรียบเทียบว่าใครมีความชำนาญในการเขียนภาพมากที่สุด หรือว่าใครมีความคิดสร้างสรรค์มากที่สุด
3. เปรียบเทียบกับเกณฑ์มาตรฐานเป็นการเปรียบเทียบที่ตั้งอยู่บนพื้นฐานของวัตถุประสงค์ในการสอน และกิจกรรม เพื่อให้ผู้เรียนได้บรรลุผลตามวัตถุประสงค์ที่ตั้งไว้

นอกจากนั้น วิรัตน์ พิชญไพบูลย์ (2520 : 63 – 69) ได้กล่าวถึงแนวทางในการประเมินผลภาคปฏิบัติไว้ดังนี้

1. ด้านทักษะในงานศิลปะ การตรวจให้คะแนน เน้นทักษะในการเขียนได้อย่างถูกต้องงดงามและเรียบร้อย ประกอบด้วย ความสามารถในการใช้แสงและเงา ความสามารถในการเขียนภาพด้วยวิธีการต่าง ๆ ความสามารถในการใช้วัสดุและเครื่องมือและความสามารถในการทำงานศิลปะได้อย่างเรียบร้อยเหมือนแบบ

2. ด้านความคิดและการออกแบบ การตรวจให้คะแนนจะเน้นการออกแบบการใช้จินตนาการต่อเติมและการจัดภาพตกแต่ง ประกอบความสามารถในการใช้หลักทางศิลปะ ความสามารถในการสร้าง และแสดงความสามารถในการสร้าง และค้นพบความสัมพันธ์ในการออกแบบ

3. ด้านการนำไปใช้การตรวจให้คะแนน เน้นการแก้ปัญหา และการนำไปใช้ได้สมเหตุสมผล เช่น ความสามารถในการแก้ปัญหาการใช้ทั่วไปและความสามารถในการมองเห็นแบบลักษณะโครงสร้างทางศิลปะที่เหมาะสมในความงามและการใช้สอย

4. ด้านความนิยมในความงาม การตรวจให้คะแนนเน้นความสามารถในการพิจารณาประเมินผล การเลือกได้ถูกต้อง เหมาะสม รวดเร็ว เช่น ความนิยมที่เกิดจากความรูสึกภายใน และความนิยมที่อาศัยประสบการณ์และการปฏิบัติ

5. ด้านความแม่นยำการตรวจให้คะแนน เน้นการให้คะแนนอยู่กับการเขียนมีส่วนสัดส่วน ช่วงจังหวะที่ถูกต้อง แน่นนอน รวดเร็ว และได้ภาพจำนวนมาก เช่น ความสามารถในการทำงานได้ถูกต้อง รวดเร็ว และความสามารถในการเข้าใจในสัดส่วน ขนาด สี และรายละเอียดของศิลปะได้แน่นอน ถูกต้องรวดเร็ว

6. ด้านการเสนองานศิลปะ การตรวจให้คะแนนเน้นการให้คะแนนที่การเขียน การจัดองค์ประกอบ และการเขียนบรรยายประกอบสั้น ๆ ได้เหมาะสม น่าเลื่อมใส เช่น ความสามารถในการสร้างงานศิลปะได้ถูกต้องตรงตามสิ่งที่ประสงค์หรือคำสั่ง และความสามารถในการอธิบาย หรือเขียนคำบรรยายงานศิลปะให้ผู้อื่นเข้าใจได้

2.11.2 เกณฑ์ในการประเมินผลศิลปะภาคปฏิบัติ

เกณฑ์ (Criteria) หมายถึง ระดับหรือมาตรฐานที่ถือว่าเป็นความสำเร็จของการดำเนินงานหรือผลของการดำเนินงาน เกณฑ์จึงเป็นตัวตัดสินคุณภาพของการปฏิบัติหรือผลที่ได้รับ เกณฑ์อาจได้มาจากมาตรฐานทางวิชาชีพ มาตรฐานการกระทำ หรือระดับความคาดหวังที่พึงประสงค์ของกลุ่มผู้เกี่ยวข้องต่าง ๆ เช่น มาตรฐานที่กำหนดโดยผู้เชี่ยวชาญ ความต้องการที่แท้จริงของผู้รับบริการ ค่านิยมของสังคม เป็นต้น (ศิริชัย กาญจนวาสี. 2536 : 76)

เราสามารถจำแนกเกณฑ์ออกเป็น 2 ประเภท คือ เกณฑ์สัมพัทธ์ (Relative criteria) และ เกณฑ์สัมบูรณ์ (Absolute criteria) โดยเกณฑ์สัมพัทธ์ เป็นเกณฑ์ที่พัฒนามาจากประสบการณ์ในการเปรียบเทียบผลระหว่างโครงการ หรือเปรียบเทียบกับผลที่ทำมาแล้ว หรือเปรียบเทียบกับปกติวิสัย (Norm) ของการจัดโครงการโดยทั่วไป สำหรับเกณฑ์สัมบูรณ์เป็นเกณฑ์ที่พัฒนามาจากหลักเหตุผลเกี่ยวกับมาตรฐานของสิ่งนั้น หรือความคิดเห็นของผู้เชี่ยวชาญอันเป็นที่ยอมรับ หรือคุณภาพของสิ่งนั้นอันเป็นที่ยอมรับของผู้เกี่ยวข้องอื่น ๆ

เมื่อพิจารณาถึงกฎเกณฑ์ที่จะนำมาใช้ตัดสินหรือประเมินคุณค่าผลงานศิลปะนั้น คลีเวอร์ (Cleaver. 1966 : 43 – 44) ได้ศึกษาปัญหาการตัดสินคุณค่าทางศิลปะ ซึ่งการตัดสินคุณค่างานศิลปะประกอบด้วยการตัดสินใจทั้งในส่วนที่มองเห็นและมองไม่เห็น บางครั้งสิ่งที่เป็นไปได้ในการพิจารณา คือ พิจารณาความชำนาญของผู้สร้างสรรค์ผลงาน ในการนำเอาความทรงจำมาสร้างให้ปรากฏ แต่ความซับซ้อนจะเกิดขึ้นในงานที่ไม่ปรากฏเป็นวัตถุที่มองเห็นได้ และงานที่เป็นนามธรรม เราสามารถตัดสินด้วยการจำกัดขอบเขตของสิ่งที่นำมาแสดง และการกำหนดขอบเขตเรื่องราว เวลาและสถานที่ ศิลปินต้องสร้างอารมณ์ความรู้สึกให้เกิดแก่ผู้ชมด้วย ทักษะในการตัดสินคุณค่ามีอยู่ด้วยกันจำนวนมาก แต่มีที่คาบเกี่ยวกันอยู่ 3 ส่วน คือ

1. ลัทธิวัตถุ (Objectivism) สมมติให้เกณฑ์การตัดสินมีความสมบูรณ์ได้มาตรฐาน ไม่สามารถเปลี่ยนแปลงได้

2. ลัทธิจิตใจ (Subjectivism) พิจารณาการตัดสินงานศิลปะจากจิตใจของแต่ละคน

3. ลัทธิความสัมพันธ์ (Relativism) การมองในแง่นี้ ถือว่าวัตถุใด ๆ ที่อยู่ในวัฒนธรรมจะมีมาตรฐานด้วย การตัดสินผลงานต้องตัดสินตามพื้นฐานทางวัฒนธรรมของผลงานนั้น ๆ ไม่ใช่ตัดสินตามวัฒนธรรมของผู้ตัดสิน

เบียร์ดสเลย์ (Beardsley, 1958 : 78) ระบุเกณฑ์กลางมาตรฐานสากลของตนในการประเมินค่าของผลงานไว้ดังนี้

1. เอกภาพ เป็นเกณฑ์ประเมินคุณค่ารวม ๆ ด้านต่าง ๆ คือ การจัดภาพ ความเป็นตัวของตัวเอง ความสัมพันธ์ของวัสดุและเทคนิค และเรื่องราว

2. ความละเอียดซับซ้อน เป็นเกณฑ์ประเมินเกี่ยวกับสิ่งต่อไปนี้ คือ การเลือกใช้องค์ประกอบที่มองเห็นได้หลากหลายไม่ซ้ำกันเกินไป และจินตนาการที่ซับซ้อน

3. ความเข้มข้นของความรู้สึก เป็นเกณฑ์ประเมินปฏิกิริยาทางความรู้สึกที่สามารถรู้สึกได้จากการพิจารณาผลงานนั้น ประกอบด้วย ความมีชีวิตชีวาไม่จืดจาง ความกล้าในการแสดงออก การกระตุ้นความสนใจ และมีแนวเรื่องที่เด่นชัด เช่น สนุกสนาน เจ็บปวด ฯลฯ

มิเชล (Michael, 1983 : 39 – 41) ได้เสนอเกณฑ์ที่ใช้ในการวัดและประเมินผลงานศิลปะภาคปฏิบัติ โดยแบ่งออกเป็นด้านต่าง ๆ ดังนี้

1. การสื่อสารข้อคิดเห็น การแสดงออก (Communication/ Expression) ได้แก่ การแสดงความสามารถด้านความคิดเห็นที่มีต่องานศิลปะ การแสดงความรู้สึกทางอารมณ์ได้อย่างชัดเจนในงานศิลปะการแสดงออกถึงความเข้าใจในเนื้อหา ทฤษฎีของงานศิลปะ การแสดงออกที่เป็นลักษณะเฉพาะตัวการแสดงออกความคิดเห็นจากการสื่อความหมายของงานศิลปะแขนงต่าง ๆ

2. ความเชื่อมั่น ความมั่นใจในตนเอง (Confidence/Self – esteem) ได้แก่ ความมั่นใจในความสามารถของตนเองด้านความคิด ความรู้สึก และความเข้าใจในการสื่อความหมายจากงานศิลปะ ความมั่นใจในการใช้สื่อของศิลปะเพื่อการประยุกต์ มีความพึงพอใจ ภาคภูมิใจในงานของตน ยอมรับต่อการเปลี่ยนแปลงความคิดร่วมกับผู้อื่น สร้างความพอใจ ความภาคภูมิใจจากผลงานของตนเอง การเน้นจุดเด่น การปรับปรุงจุดด้อยของตนเอง ความสามารถในการตัดสินใจ การเป็นผู้นำในการแสดงออกทางศิลปะและความสามารถในการสร้างความสัมพันธ์ระหว่างคุณค่าทางสุนทรียภาพกับชีวิตของตน

3. การพัฒนาการรับรู้ (Perceptual development) ได้แก่ พัฒนาการรับรู้ส่วนประกอบพื้นฐานศิลปะด้วยจักขุสัมผัส พัฒนาการรับรู้ส่วนประกอบพื้นฐานศิลปะด้วยโสตสัมผัส ฯลฯ มีความเข้าใจถึงคุณค่าการมองแบบจักขุสัมผัส โสตสัมผัส ฯลฯ ที่แตกต่างกันและมีความรู้สึกถึงคุณค่าของรูปร่าง สี เส้น น้ำหนัก องค์ประกอบของศิลปะ

4. องค์ความรู้ทางสุนทรียภาพและความต่อเนื่อง (Aesthetic organization and consistency) ได้แก่ มีความรู้ในการสร้างความสัมพันธ์ระหว่างความคิดและวัตถุดิบเพื่อการแสดงออก มีระบบประสาทในการจัดองค์ประกอบศิลปะ มีความเข้าใจในคุณค่าของความคิด ความรู้สึกความเข้าใจหลักการจัดภาพเป็นความสมดุล จังหวะ ความแตกต่าง ความประสานกลมกลืน ความขัดแย้ง ฯลฯ มีความสามารถในการจัดภาพให้มีเอกภาพ มีเจตคติที่ดีต่อสิ่งแวดล้อม มีความสามารถในการแก้ไขปัญหาด้านสุนทรียภาพและมีทัศนคติที่ดีต่อสุนทรียภาพในรูปแบบต่าง ๆ

5. ความรู้สึกลงในสาขาศิลปะ (Knowledge of the field of art) ได้แก่ ความรู้ ความเข้าใจในเชิงปฏิบัติการการใช้เครื่องมือ มีจินตนาการในการใช้เครื่องมือ มีความรู้ ความเข้าใจในการวิเคราะห์รูปแบบโครงสร้างของงานศิลปะ มีความรู้ ความเข้าใจในการประเมินค่าการทำงานสื่อต่าง ๆ และการแสดงความคิดเห็นงานศิลปะทั้งในอดีตและปัจจุบัน มีความรู้ ความเข้าใจในศิลปะนามธรรม และมีความรู้ ความเข้าใจในรูปแบบศิลปะต่าง ๆ

6. ความคิดริเริ่มสร้างสรรค์ (Creativity) ประกอบด้วย มีความรู้สึกตระหนักถึงปัญหาในการทำงานศิลปะ มีความคิดเฉพาะเป็นของตนเอง แสดงถึงความคิดสร้างสรรค์ ในการอธิบายความคิด ความรู้สึกของตนเอง นำเส้น รูปทรง สี แบบ วัตถุต่าง ๆ มาสร้างเป็นสิ่งแปลกอย่างน่าทึ่ง รู้จักสำรวจแนวทางในการแก้ไข ปรับปรุงผลงาน การพัฒนาคุณลักษณะพิเศษของตนเอง มีทักษะในการคิดมีวิธีปรับปรุงผลงานด้วยวิธีการใหม่ ๆ ความสามารถวิเคราะห์งานศิลปะจากส่วนรวมไปส่วนย่อยและความสามารถในการจัดองค์ประกอบ

7. ทักษะความชำนาญ (Skill/Craftsmanship) ได้แก่ การพัฒนาการใช้มือ การควบคุมการทำงาน ทักษะในการใช้เส้น รูปทรง สี ในการสร้างสรรค์และทักษะในการใช้เครื่องมือรักษาเครื่องมือ

8. ความเพลิดเพลิน ความพึงพอใจ (Enjoyment/Satisfaction) ได้แก่ ความพึงพอใจในการทำงาน การหาโอกาสในการทำงาน การกระตุ้นผู้อื่นเข้าร่วมประสบการณ์ในการทำงานด้วย ความพึงพอใจในผลงาน และประสบการณ์ ความต้องการในการอธิบายงาน การสนทนาศิลปะกับผู้อื่นและการมีส่วนร่วมในประสบการณ์ศิลปะแขนงต่าง ๆ ผลิตภัณฑ์ใหม่ ๆ และการฟังแนวความคิดเห็น

ไอสนอร์ (Eisner. 1972 : 214 – 216) กล่าวถึงเกณฑ์การประเมินผลงานศิลปะไว้ดังนี้

1. ความชำนาญทางเทคนิคที่ปรากฏในผลงาน ได้แก่ ความสามารถในการใช้และควบคุมอุปกรณ์ว่าผู้เรียนสามารถควบคุมอุปกรณ์เหล่านี้ได้มากน้อยเพียงใด

2. สุนทรียภาพและลักษณะการแสดงออกของผลงาน ได้แก่ ความสามารถในการจัดรูปทรงผลงาน การแสดงออก รวมถึงความสัมพันธ์ระหว่างเทคนิคและสุนทรียภาพ

3. จินตนาการและการสร้างสรรค์ในผลงาน ได้แก่ การประเมินระดับของความคิด สร้างสรรค์และจินตนาการที่ปรากฏ ความเฉลียวฉลาด ความแปลกใหม่และความรู้สึก

จากข้อมูลดังกล่าวข้างต้นสรุปได้ว่า สภาพปัจจุบันการประเมินผลงานศิลปะในภาคปฏิบัติของครูผู้สอนมักมีเกณฑ์ของตนเอง และคำนึงถึงความรู้สึกของตนเองมากกว่าการมีหลักเกณฑ์ที่ให้คะแนนอย่างถูกต้อง ดังนั้นเพื่อให้ผู้สอนมีการประเมินผลงานศิลปะอย่างมีการหลักเกณฑ์การประเมินที่ถูกต้องและชัดเจนและได้ผลงานที่มีคุณค่า มีแนวทางที่เป็นกลาง เพื่อการพัฒนาการประเมินผลงานศิลปะให้มีประสิทธิภาพยิ่งขึ้น

2.11.3 โครงสร้างของภาพ

โครงสร้างของภาพ หรือนักวิชาการบางท่านเรียกว่าส่วนประกอบศิลป์ เป็นส่วนประกอบส่วนหนึ่งในการสร้างสรรค์ศิลปะที่จะช่วยให้การนำเสนอแนวความคิดของศิลปินเป็นไปตามวัตถุประสงค์ อารี สุทธิพันธุ์ กล่าวไว้ว่า ทางทัศนศิลป์เขาแบ่งส่วนประกอบ (Element) ไว้หลายชนิดด้วยกันคือเส้น (Line) รูปทรง (Shape Form) ทิศทาง ตำแหน่ง แรง (Direction Position, Forces) ขนาด ระยะทาง (Size & Tone) น้ำหนัก (Size & Distance) ลักษณะผิว (Texture & Surface)สี (Color)

วิรุณ ตั้งเจริญ ได้แสดงว่าทักษะว่า ในสมัยปัจจุบันมักจะใช้คำว่าโครงสร้างแทนคำว่า องค์ประกอบศิลป์ ซึ่งจะมีความหมายกว้างกว่า อันหมายถึงทุกอย่างที่ศิลปินนำมาใช้ในการสร้างภาพไม่ว่าจะเป็นมิติ ระบาย จังหวะหรือ ความกลมกลืน

สุชาติ เกาทอง กล่าวถึงโครงสร้างทางทัศนศิลป์ว่าโครงสร้างทางทัศนศิลป์ หรือบางทีภาษาทางศิลปะเรียกว่าเป็นองค์ประกอบหลักในการสร้างสรรค์และพิจารณางานศิลปะว่ามีความงามและความเหมาะสมมากน้อยเพียงใด โครงสร้างทางทัศนศิลป์ประกอบด้วยดุลยภาพ (balance) สัดส่วน (proportion) ความกลมกลืน (harmony) ช่วงจังหวะ (rhythm) และจุดเด่นของงานศิลปะ (dominance) แต่หลักการเหล่านี้ไม่ได้เป็นกฎเกณฑ์ตายตัวที่จะต้องปฏิบัติตามทุกกรณีเพราะการจัดและการสร้างสรรค์เป็นการแก้ปัญหาซึ่งอาจต้องดัดแปลงปรับปรุงเพื่อให้งานนั้นมีเอกภาพมากที่สุด

เทียนชัย ตั้งพรประเสริฐ ได้อธิบายโครงสร้างทางทัศนศิลป์ที่สุชาติ เกาทองกล่าวถึงว่าเป็นหลักขององค์ประกอบศิลป์อันประกอบด้วย

เอกภาพ (Unity) ในงานทัศนศิลป์ทุกสาขาเอกภาพมีความสำคัญและมีความจำเป็นต้องสร้างให้เกิดขึ้นเป็นเอกภาพของเนื้อหา เรื่องราว รูปแบบความคิดและเอกภาพของส่วนประกอบสำคัญของศิลปะ หรือรูปร่าง รูปทรงก็ตาม

ดุลยภาพ (Balance) หมายถึงการจัดส่วนประกอบย่อยต่าง ๆ ได้แก่ จุด เส้น สี น้ำหนัก พื้นผิว บริเวณว่างหรือรูปร่าง รูปทรง ภายในผลงานศิลปะให้เกิดความสมดุลหรือเท่าเทียมกันโดยมีเส้นแกนสมมติ 2 เส้นเป็นตัวกำหนดดุลยภาพ เพื่อให้รูปร่างรูปทรงที่ปรากฏในภาพเกิดความสมดุลในลักษณะใดลักษณะหนึ่ง เช่น แบบซ้ายขวาเหมือนกันหรือภาพซ้ายขวาไม่เหมือนกัน

จุดเด่นและการเน้น (Dominance & Emphasis) หมายถึง ส่วนที่สำคัญในการมีความชัดเจน สะดุดตาเป็นแห่งแรก รับรู้ได้ด้วยการมองผลงานที่สำเร็จแล้ว มีลักษณะมีอำนาจ ตระหง่าน เด่นชัดกว่าส่วนประกอบอื่น ๆ

สุชาติ สุทธิ กล่าวไว้ในหนังสือเรียนรู้การเห็น พื้นฐานการวิจารณ์ศิลปะว่า การเห็นการสื่อความหมายนอกจากจะเป็นขั้นต้นของผู้ดูหรือรับรู้ศิลปะ (art perceiver) จากผลงานของผู้สร้างศิลปะ (art creator) แล้วยังเป็นประโยชน์ต่อการสร้างสรรค์ศิลปะจะได้นำไปใช้เพื่อถ่ายทอดการแสดงออก (expression) ของเขาออกมาได้เป็นอย่างดีภายใต้หลักการอันเดียวกัน หรือกล่าวโดยสรุปอีกนัยหนึ่งก็คือการเห็นและสื่อความหมาย นอกจากใช้เป็นเครื่องมือในการเรียนรู้ของผู้ดูในขณะที่กำลังเผชิญหน้ากับผลงานศิลปะแล้ว ยังเป็นประโยชน์สำหรับผู้สร้างสรรค์ศิลปะที่จะใช้เป็นสื่อ เพื่อถ่ายทอดความรู้สึกของตนเองลงสู่ผลงานศิลปะในช่วงขณะที่ทำการสร้างสรรค์อีกด้วย และที่มาของตัวสื่อที่จะให้ความหมายในการเห็นก็คือ ค่าของงานความสัมพันธ์ที่เกิดขึ้นระหว่างมูลฐานทัศนศิลป์ (Visual elements) กับหลักการทัศนศิลป์ (Visual principle) ผู้ที่สร้างสรรค์หรือศิลปินได้นำมาใช้สร้างให้เป็นตัวผลงานออกมาและมีคุณค่าทางสุนทรียภาพทั้งความงามและการแสดงออกที่สามารถรับรู้และเรียนรู้ได้

จากมูลฐานทัศนศิลป์และหลักการทัศนศิลป์ที่สุชาติ สุทธิ กล่าวถึง พอสรุปเป็นหัวข้อ ดังนี้มูลฐานทัศนศิลป์ได้แก่ เส้น รูปร่างรูปทรง แสงและสี บริเวณว่างและมิติ ผิว

หลักการทัศนศิลป์ได้แก่ ดุลยภาพ จังหวะและความเคลื่อนไหว การลดหลั่น ความกลมกลืนการย่ำและการตัดกันและความหลากหลายและสัดส่วน

จากองค์ประกอบดังกล่าว ผู้วิจัยได้เลือกหลักการพื้นฐานมาใช้ในการประเมินผลงานศิลปะของนักเรียน

ชั้นประถมศึกษาปีที่ 6 ของโรงเรียนสุเหร่าคลองหนึ่งในด้านต่าง ๆ ดังต่อไปนี้

ด้านเนื้อหาประกอบด้วย การเน้นจุดเด่นของภาพ ความสมดุลของภาพ การระบายสี การใช้สี ความคิดสร้างสรรค์ ความประณีต ความสะอาด รวมทั้งหมด 7 ประเด็นการประเมินผลงานศิลปะของนักเรียน

2.11.4 วิธีการและเครื่องมือในการประเมินผลวิชาศิลปะ

อุบล ตูจินดา (2533 : 219 – 222) กล่าวว่าไว้ว่า การประเมินผลความเจริญงอกงามทางศิลปะของนักเรียน มีเทคนิคและวิธีต่าง ๆ หลายประการ เช่น

1. การประเมินผลตนเอง (Self Evaluation) นักเรียนอาจได้รับคำแนะนำให้ประเมินตัวเองได้หลายวิธี เช่น สังเกตผลงานของตนเองที่ทำไปแล้ว กับผลงานชิ้นใหม่หรืออาจทำเป็นตารางและให้คะแนนตนเอง การประเมินผลตนเองจะช่วยให้นักเรียนรู้จักควบคุม มีวินัยในตนเอง และรู้จักพิจารณาไตร่ตรองในผลงาน

2. การประเมินผลของครู (Teacher Evaluation) ครูอาจจะใช้วิธีการหลาย ๆ อย่าง เพื่อรวบรวมข้อมูลความก้าวหน้าของนักเรียน เช่น การสังเกต การวัดผลงานและการใช้แบบทดสอบ ฯลฯ

3. กระบวนการสังเกต (Observation Procedures) ครูอาจจะรวบรวมจากพฤติกรรมของนักเรียนดังนี้

- สังเกตพฤติกรรมขณะการสอน เป็นการสังเกตการทำงานของนักเรียน ในแง่มุมต่าง ๆ เช่น การเลือกและการใช้เครื่องมือ การเอาใจใส่ การตั้งใจทำงาน การร่วมมือกันทำงาน การควบคุมอารมณ์ - การสังเกตพฤติกรรมหลังการสอน เป็นการสังเกตการแสดงออกและความเป็นไปของนักเรียนหลังจากเรียนแล้ว เช่น เมื่อนักเรียนเรียนเรื่องเทคนิคการใช้สีแล้ว ครูคอยสังเกตดูว่า นักเรียนรู้จักวิธีการใช้สีหรือไม่ การสังเกตแบบนี้ ควรมีการบันทึกโดยทำตาราง การวัดแบบสังเกตไว้ด้วย นอกจากการสังเกตพฤติกรรมแล้ว ครูควรจะสังเกตรายละเอียดของการทำงานและสังเกตคุณภาพของงานอีกด้วย

4. ระเบียบสะสม (Anecdotal Records) ครูอาจใช้แผ่นกระดาษ วางไว้บนโต๊ะ สำหรับที่จะจดบันทึกได้ทันที กระดาษนี้ แบ่งออกเป็นสองช่อง ช่องหนึ่ง สำหรับบันทึกพฤติกรรมที่มองเห็นได้ อีกช่องหนึ่ง เป็นการตีความหมายของพฤติกรรม เช่น

สิ่งที่เกิดขึ้น : จากการอธิบายเรื่องการพิมพ์โดยใช้วัสดุต่าง ๆ จิตวิทยาได้เสนอแนะวิธีการพิมพ์ด้วยเปลือกหอยที่มีลวดลายต่าง ๆ

การแปลความ : แสดงของจุดสนใจของจิตวิทยาอยู่ที่เรื่องนี้และนำไปสัมพันธ์กับเปลือกหอยชนิดต่าง ๆ ที่พบที่ชายหาด

การฝึกฝนวิธีการเช่นนี้อย่างสม่ำเสมอ ทำให้ครูสามารถใช้ประโยชน์ต่อการประเมินผลพัฒนาการของนักเรียนได้ด้วย

5. ตารางตรวจสอบรายการ (Checklists) ฐานที่เป็นครู ต้องการจะทราบว่านักเรียนมีพฤติกรรมนั้น ๆ หรือไม่ เพราะบางครั้งการสังเกตโดยไม่มีรายการให้ตรวจสอบ ครูก็อาจจะลืมสังเกตพฤติกรรมหรือบางครั้ง เมื่อครูมีปัญหหรือสงสัยในกระบวนการเรียนการสอนครูสามารถค้นหาได้ว่า นักเรียนคนใดบ้างที่สามารถแสดงพฤติกรรมอะไร เพราะเหตุใด การใช้ตารางตรวจสอบ (Check Lists) จะช่วยประกอบการพิจารณาของครูได้ เพราะคำถามต่าง ๆ จะนำไปสู่การตั้งเกณฑ์การให้คะแนน เช่น ในแบบตรวจสอบรายการหนึ่ง มีพฤติกรรมต้องการสังเกต ครูจึงกำหนดคะแนนไว้แต่ละข้อ และเมื่อรวมคะแนนทั้งหมดแล้ว จึงใช้เป็นเครื่องมือตัดสินคะแนนครั้งสุดท้าย

6. การจัดอันดับคุณภาพ (Rating Scales) เป็นการให้นั้หนักพฤติกรรมที่สังเกตได้ การจัดอันดับคุณภาพคือ ลักษณะพฤติกรรมที่จะวัด ต้องตรงกับจุดประสงค์ เป็นลักษณะที่ชัดเจน สามารถสังเกตได้ วัดได้ และเมื่อวัดผลงานของนักเรียนหลาย ๆ คนแล้ว ใช้วิธีการจัดอันดับคุณภาพของงานนี้

- นำงานมาพิจารณาอย่างละเอียดถี่ถ้วน แล้วแบ่งงานออกเป็น 3 พวก คือ ดี ปานกลางและด้อย
- นำงานแต่ละกลุ่มมาพิจารณา วัด เรียงลำดับ จากคุณภาพดีไปหาด้อย
- นำอันดับแต่ละกลุ่มมาเรียงติดต่อกัน จากดีไปหาด้อยที่สุด อาจจะสลับเปลี่ยนตามความเหมาะสมได้อีกแล้ว

ตรวจดูความสมบูรณ์อีกครั้งหนึ่ง

- แล้วนำมาคิดเป็นเกรด สำหรับชั้นในชั้นประถมศึกษา ให้วัดผลเพียงผ่านไม่ผ่านเท่านั้น

7. การทดสอบ (Testing) เป็นการทดสอบในการประเมินผลตามจุดประสงค์ต่าง ๆ เช่น การวัดความรู้ วัดผลสัมฤทธิ์ทางการเรียน ความถนัด เซาว์และบุคลิกภาพ แบ่งออกได้ดังนี้

- การสอบภาคปฏิบัติคือ การวัดวิธีปฏิบัติงาน (Procedure) และการวัดผลงาน (Product) เช่น

1. วัดว่านักเรียน ใช้วัสดุและเครื่องมือถูกต้องหรือไม่ การสอบแบบนี้ จะใช้วิธีให้นักเรียนบอกและหยิบวัสดุ และเครื่องมือวางลำดับการใช้ให้ถูกต้องว่า ควรใช้เครื่องมือชนิดใด ทำอะไร

2. ครูกำหนดงานให้นักเรียนทำ แล้วครูสังเกตวิธีการ หรือกระบวนการว่านักเรียนทำถูกหรือไม่ แล้วพิจารณาให้คะแนนตามคุณภาพของงานนั้น ๆ

- 3. เมื่อนักเรียนได้ปฏิบัติงานนั้นแล้ว ครูตรวจผลงานตามจุดประสงค์ที่ตั้งไว้

จากเครื่องมือในการประเมินผลวิชาศิลปะ ผู้วิจัยได้ใช้การประเมินแบบมาตรวัดประเมินค่า (Rating Scales) มาใช้ในการประเมินผลงานศิลปะในครั้งนี้

3. เอกสารและงานวิจัยที่เกี่ยวข้อง

1. งานวิจัยในต่างประเทศ

งานวิจัยเกี่ยวกับการวิเคราะห์ค่าความเชื่อมั่นของผู้ประเมินที่ผู้วิจัยได้ค้นคว้ามามีดังนี้

1.1 งานวิจัยที่เกี่ยวข้องกับความเชื่อมั่นของผู้ประเมิน (Rater Reliability)

เทสส์ และคนอื่น ๆ (Tasse and others. 1994) ศึกษาวิจัยเกี่ยวกับ ความสอดคล้องของผู้ประเมิน เรื่อง แบบทดสอบซ้ำ และความเชื่อมั่นระหว่างผู้ประเมินแบบวัดพฤติกรรมการปรับตัวของ คิวเบก (Quebec) ซึ่งได้รับการนำไปใช้อย่างแพร่หลายในเมืองคิวเบก (ประเทศแคนาดา) เพื่อประเมินพฤติกรรมของผู้ที่มีระดับสติปัญญาต่ำในสภาพแวดล้อมเกี่ยวกับการศึกษา การทำงาน การอยู่อาศัย หรือ สภาพแวดล้อมในโรงพยาบาล การศึกษานี้ประเมินความเชื่อมั่นของผู้ตรวจและความเชื่อมั่นโดยวิธีสอบซ้ำของ QABS และเพื่อวัดความเชื่อมั่นของการสอบซ้ำ QABS จึงได้รับการประเมินโดยผู้ตรวจจำนวน 27 คน กับกลุ่มตัวอย่างจำนวน 45 คน (อายุระหว่าง 6 ถึง 47 ปี) จากบ้านเรือน สถานที่พักอาศัย โรงเรียน และโรงพยาบาลต่างๆ และเพื่อให้ได้ค่าความเชื่อมั่นของผู้ตรวจ ได้ใช้กลุ่มตัวอย่างอีกกลุ่มจำนวน 44 คน (อายุระหว่าง 3 ถึง 54 ปี) แต่ละคนจะได้รับการประเมินจากผู้ตรวจ 2 คน ค่าสัมประสิทธิ์ความเชื่อมั่นระหว่างการตรวจโดยใช้ประเมินภาพรวม (Global Score) กับระดับพฤติกรรมปรับตัว มีค่าเท่ากับ .92 คำนวณโดยวิธีสอบซ้ำ และระหว่างผลรวมของคะแนนดิบ มีค่าเท่ากับ

.99 ค่าสัมประสิทธิ์ความสอดคล้องของผู้ตรวจที่ใช้การตรวจแบบประเมินภาพรวม (Global Score) กับระดับพฤติกรรมกำหนัด มีค่าเท่ากับ .83 และ ระหว่างผลรวมของคะแนนดิบ เท่ากับ .99 พบว่า QABS เป็นมาตรวัดพฤติกรรมกำหนัดแบบอิงเกณฑ์ที่มีความเชื่อถือได้สูง มีเกณฑ์ปกติ และแบบเกณฑ์การให้คะแนนที่ชัดเจนง่ายต่อการใช้งาน

ไอโซนิโอ (Isonio. 1991) ศึกษาวิจัยเกี่ยวกับความเชื่อมั่นของผู้ตรวจ เรื่องผู้ตัดสินตัวอย่างงานเขียนของวิทยาลัยโกลเดนเวสต์ : การประเมินความเชื่อมั่นของผู้ตรวจ งานวิจัยนี้ต้องการศึกษาตัวอย่างงานเขียนของนักเรียน โดยให้คะแนนแบบประเมินภาพรวม (Holisticly) โดยใช้ผู้ตรวจเป็นรายคู่ ด้วยเกณฑ์การวัดแบบ 6 ระดับ (Six-point Scale) ผู้ตรวจต้องทำเครื่องหมาย + หรือ - วิธีนี้เป็นการเปลี่ยนจากการให้คะแนนแบบจำนวนเต็มมาเป็นทศนิยม ในปี 1991 ตัวอย่างงานเขียนจำนวน 499 ตัวอย่าง ที่นำจาก ส่วนหนึ่งของการทดสอบเพื่อจัดอันดับสำหรับนักเรียนในโครงการ SOAR ของ Coast Community College District (เขตวิทยาลัยชุมชนแถบชายฝั่ง) ที่วิเคราะห์การให้คะแนนที่น่าเชื่อถือของผู้ตรวจ งานวิจัยนี้ใช้ 2 วิธี เพื่อประเมินความคงที่ในการตัดสินตัวอย่างงานเขียน วิธีแรกที่ใช้เป็นการคำนวณความแตกต่างระหว่างค่าบ่งบอกระดับตัวอย่างเขียน 2 ค่า มีค่าเท่ากับ 34.4 % ของคู่ที่ตัดสินทั้งหมด ค่าบ่งบอกระดับจากผู้ตัดสินแต่ละคนจะเหมือนกัน แต่ละค่าจะแตกต่างกันไป 1 ใน 3 ระดับ มี 30.9 % ของตัวอย่าง และ 2 ใน 3 ระดับ มี 15.4 % ของตัวอย่าง และ 1 ระดับ มี 13.4 % โดยรวม แล้วค่าบ่งชี้ระดับจะแตกต่างออกไป 1 ระดับ หรือน้อยกว่า 94.1 % ของตัวอย่าง ซึ่งเกินมาตรฐานต่ำสุดที่ 90 % ที่แนะนำโดย California Community Colleges (CCC) วิธีที่สองที่ใช้ก็คือ สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน แสดงความสัมพันธ์ระหว่างการตรวจเป็นคู่ สำหรับตัวอย่างข้อเขียน ค่าความสัมพันธ์มีค่าเป็นบวก ($r = .76$) ซึ่งเกินมาตรฐานต่ำสุดที่ .75 ตามคำแนะนำของ CCC และข้อมูลที่ได้อบรมรับสมมติฐานที่ว่า การตรวจจากผู้ตรวจแต่ละคนได้ผลสอดคล้องกัน

สตีเฟน (Steven. 1991) ได้ศึกษาเกี่ยวกับความเชื่อมั่นของผู้ตรวจ เรื่อง ผู้ตัดสินตัวอย่างงานเขียนของวิทยาลัยโกลเดนเวสต์ : การประเมินความเชื่อมั่นของผู้ตรวจ จุดประสงค์ต้องการศึกษาตัวอย่างงานเขียนของนักเรียน โดยให้คะแนนแบบประเมินภาพรวม (Holisticly) โดยใช้ผู้ตรวจเป็นรายคู่ ด้วยเกณฑ์การวัดแบบ 6 ระดับ (Six-point Scale) ผู้ตรวจต้องทำเครื่องหมาย + หรือ - วิธีนี้เป็นการเปลี่ยนจากการให้คะแนนแบบจำนวนเต็มมาเป็นแบบทศนิยม ในปี 1991 ตัวอย่างงานเขียนจำนวน 499 เรื่อง ที่นำมาจากส่วนหนึ่งของการทดสอบเพื่อจัดอันดับสำหรับนักเรียนในโครงการ SOAR ของ Coast Community College District (เขตวิทยาลัยชุมชนแถบชายฝั่ง) ที่วิเคราะห์การให้คะแนนที่น่าเชื่อถือของผู้ตรวจ งานวิจัยนี้ใช้ 2 วิธี เพื่อประเมินความคงที่ในการตัดสินตัวอย่างงานเขียน วิธีแรกที่ใช้เป็นการคำนวณความแตกต่างระหว่างค่าบ่งบอกระดับตัวอย่างงานเขียน 2 ค่า มีค่าเท่ากับ 34.4 % ของคู่ที่ตัดสินทั้งหมด ค่าที่บ่งบอกระดับจากผู้ตัดสินแต่ละคนจะเหมือนกัน แต่ละค่าจะแตกต่างกันไป 1 ใน 3 ระดับ มี 30.9 % ของตัวอย่างและ 2 ใน 3 ระดับมี 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 และข้อมูลที่ได้อบรมรับสมมติฐานที่ว่า การตรวจจากผู้ตรวจแต่ละคนได้ผลสอดคล้องกัน

2. งานวิจัยภายในประเทศ

งานวิจัยภายในประเทศเกี่ยวกับความเชื่อมั่นของผู้ตรวจให้คะแนนที่ผู้วิจัยรวบรวมได้ มีดังนี้

พรรณี เจียมสุบุตร (2543) ได้ศึกษาเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจและวิธีการตรวจต่างกัน การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจ 2 คน และ 3 คน โดยใช้วิธีการตรวจวิธีประเมินรวม และวิธีวิเคราะห์ย่อย โดยศึกษาในกลุ่มตัวอย่างเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ภาคเรียนที่ 2 ปีการศึกษา 2542 ของโรงเรียนในสังกัดกรมสามัญศึกษา จังหวัดกรุงเทพมหานคร กลุ่ม 7 จำนวน 70 คน ซึ่งเลือกมาโดยการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการศึกษาเป็นแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียน วิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 2 เรื่องสมการและอสมการ ผลการศึกษา ได้ผลดังนี้

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

2. ค่าความเชื่อมั่นของแบบทดสอบวัดความสามารถในการแก้โจทย์ปัญหาทางคณิตศาสตร์ ที่ใช้วิธีการตรวจให้คะแนนแบบวิธีรวม และวิธีวิเคราะห์ของผู้ตรวจให้คะแนน 2 คน มีค่าเท่ากับ 0.97 และ 0.92 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ส่วนแบบทดสอบที่ใช้วิธีการตรวจให้คะแนนวิธีประเมินรวม และวิธีการวิเคราะห์ ของผู้ตรวจให้คะแนน 3 คน มีค่าเท่ากับ 0.98 และ 0.96 ตามลำดับ ซึ่งแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05

ไพรัตน์ วงษ์นาม (2533 : 105 – 123) ได้ศึกษาเปรียบเทียบผลของตรวจ การชี้แจงและการรู้ผลการเรียนของผู้ตอบ ที่มีต่อสัมประสิทธิ์การอ้างอิงสรุป (Generalizability Coefficient) ของการสอบด้วยแบบทดสอบความเรียง ที่วัดความสามารถในการแสดงความคิดเห็นเกี่ยวกับข่าว และเหตุการณ์ ของนักเรียนชั้นประถมศึกษาปีที่ 6 จำนวน 30 คน ข้อสอบความเรียง 5 ข้อ และครูผู้ตรวจให้คะแนน จำนวน 20 คน สุ่มแบ่งผู้ตรวจให้คะแนนเป็น 2 กลุ่ม ๆ ละ 10 คน สำหรับการตรวจ 2 วิธี คือวิธีการประเมินรวม และวิธีวิเคราะห์ในแต่ละวิธีทำการตรวจ 3 ครั้ง คือ ครั้งแรก ตรวจโดยใช้ประสบการณ์เดิมของผู้ตรวจให้คะแนน ครั้งที่สอง ตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยโดยไม่รู้ผลการเรียนของผู้ตอบ ครั้งที่สาม ตรวจตามเกณฑ์ที่ได้รับการชี้แจงจากผู้วิจัยเหมือนครั้งที่สอง แต่รู้ผลการเรียนของผู้ตอบด้วย พบว่า เมื่ออ้างอิงสรุปไปยังเอกภาพของข้อสอบและผู้ตรวจให้คะแนนพร้อมกัน ค่าสัมประสิทธิ์การอ้างอิงสรุป (ความเชื่อมั่น) ของแบบสอบความเรียงที่ตรวจโดยวิธีประเมินรวม มีค่าอยู่ระหว่าง 0.3328 – 0.4782 และวิธีวิเคราะห์ มีค่าอยู่ระหว่าง 0.3348 – 0.5895 และใช้ข้อสอบ 5 ข้อ ใช้การตรวจวิธีวิเคราะห์ ตามที่ได้รับการชี้แจงและไม่รู้ผลการเรียนของผู้ตอบ จึงจะให้ค่าสัมประสิทธิ์การอ้างอิงสรุปไม่น้อยกว่า 0.50 เมื่ออ้างอิงสรุปไปยังเอกภาพของผู้ตรวจให้คะแนนอย่างเดียว ค่าสัมประสิทธิ์การสรุปอ้างอิง (ความเชื่อมั่น) ที่ตรวจโดยวิธีประเมินรวม มีค่าอยู่ระหว่าง 0.4743 – 0.6865 และวิธีวิเคราะห์ มีค่าอยู่ระหว่าง 0.5985 – 0.7761 และใช้ข้อสอบ 5 ข้อ ผู้ตรวจให้คะแนน 5 คน ในการตรวจวิธีวิเคราะห์ตามที่ได้รับการชี้แจง และไม่รู้ผลการเรียนของผู้ตอบ จึงจะให้ค่าสัมประสิทธิ์การสรุปอ้างอิง มีค่าไม่น้อยกว่า 0.50 เมื่ออ้างอิงไปยังเอกภาพของผู้ตรวจให้คะแนนอย่างเดียว พบว่า เฉพาะกรณีให้ผู้ตรวจให้คะแนนได้รับการชี้แจงและไม่รู้ผลการเรียนของผู้ตรวจให้คะแนนเท่านั้น ที่ให้ค่าสัมประสิทธิ์การ

สรุปอ้างอิง สูงกว่าการตรวจโดยใช้ประสบการณ์เดิม ถ้ารู้ผลการเรียนของผู้ตอบค่าสัมประสิทธิ์การอ้างอิงจะไม่แตกต่างกัน

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

จิราพร อัครสมพงศ์ (2536) ได้ศึกษาเปรียบเทียบความสอดคล้องของคะแนนแบบสอบความเรียงจากผู้ตรวจให้คะแนนและวิธีการตรวจต่างกัน ที่มีวัตถุประสงค์เพื่อ ศึกษาเปรียบเทียบความสอดคล้องในการให้คะแนนของผู้ตรวจให้คะแนน 2 คน ที่ตรวจแบบสอบความเรียงโดยใช้วิธีเดียวกัน และความสอดคล้องในการให้คะแนนระหว่างวิธีการตรวจที่แตกต่าง พบว่า ความสัมพันธ์ในการให้คะแนนระหว่างผู้ตรวจให้คะแนนด้วยวิธีเดียวกัน ผู้ตรวจให้คะแนนที่ตรวจด้วยวิธีตรวจด้วยวิธีเดียวกันมีความสัมพันธ์กันสอดคล้องกัน ในทิศทางเดียวกันและมีค่าสัมประสิทธิ์สหสัมพันธ์ของการตรวจแบบวิธีวิเคราะห์ให้เสร็จเป็นรายข้อ กับการตรวจแบบวิธีประเมินรวมให้เสร็จเป็นรายข้อมีค่าสูงสุด คือ .8116

เกศเกล้า หมุ่มมิ่ง (2539) ได้ศึกษาสูตรหาค่าความเชื่อมั่น จำนวน 6 สูตร ในการประมาณค่าความเชื่อมั่นของมาตรประเมินวัดแบบประเมินค่าด้านมโนภาพแห่งตน ซึ่งได้แก่สูตร α , r_{L1} , r_{L2} , α_x , Ω และ Ω_w โดยมีจุดมุ่งหมายเพื่อ 1) หาค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่าด้านมโนภาพแห่งตน 2) หาค่าความลำเอียงทางสถิติของการประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่าด้านมโนภาพแห่งตน 3) เปรียบเทียบค่าความเชื่อมั่นแต่ละสูตรของมาตรวัดแบบประเมินค่าด้านมโนภาพแห่งตน

กลุ่มตัวอย่างที่ใช้ในการศึกษาเป็นนักเรียนชั้นประถมศึกษาปีที่ 6 ภาคเรียนที่ 2 ปีการศึกษา 2538 ของโรงเรียนอัครสังฆมณฑลกรุงเทพฯ ที่ตั้งอยู่ในกรุงเทพมหานครจำนวน 600 คน โดยวิธีการสุ่มแบบแบ่งชั้น

ผลการวิจัยพบว่า

- 1) ความคลาดเคลื่อนมาตรฐานของการประมาณค่าความเชื่อมั่นด้วยสูตร α มีค่าต่ำสุดและสูตร r_{L1} มีค่าสูงสุด
- 2) ความลำเอียงทางสถิติของการประมาณค่าความเชื่อมั่นที่คำนวณด้วยสูตร r_{L1} มีค่าต่ำสุดและสูตร α มีค่าสูงสุด
- 3) การเปรียบเทียบค่าความเชื่อมั่นและค่าความเชื่อมั่นเฉลี่ยที่คำนวณด้วยสูตรทั้ง 6 สูตร จากกลุ่มตัวอย่าง 600 คน และกลุ่มตัวอย่างขนาด 400 คน จำนวน 200 กลุ่ม พบว่า ค่าความเชื่อมั่นที่คำนวณด้วยสูตร α มีค่าต่ำสุด และสูตร Ω_w มีค่าสูงสุด เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่นโดยใช้สูตร UX_1 พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 จากนั้นทดสอบค่าความเชื่อมั่นรายคู่ด้วยวิธีการของมาราคูโยล พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .001 ทุกคู่ ส่วนค่าความเชื่อมั่นเฉลี่ยที่คำนวณด้วยสูตรทั้ง 6 สูตร พบว่า ค่าความเชื่อมั่นเฉลี่ยที่คำนวณด้วยสูตร α มีค่าต่ำสุดและสูตร Ω_w มีค่าสูงสุด ทดสอบความแตกต่างของค่าความเชื่อมั่นเฉลี่ย โดยใช้

สูตรการวิเคราะห์ความแปรปรวนแบบตัวแปรพหุทางเดียว พบว่า มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

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 ส่วนการตรวจให้คะแนนแบบประเมินรวมโดยใช้เกณฑ์การให้คะแนน เมื่อมีจำนวนผู้ตรวจให้คะแนน 3, 4, 5 และ 6 มีคะแนนเฉลี่ยอยู่ระหว่าง 9.82 – 10.05 โดยการตรวจให้คะแนน เมื่อมีจำนวนผู้ตรวจให้คะแนน 4 คน คะแนนเฉลี่ยมีแนวโน้มสูงที่สุด รองลงมาคือ จำนวนผู้ตรวจให้คะแนน 3 และ 6 คน และ เมื่อมีจำนวนผู้ตรวจให้คะแนน 5 คะแนนเฉลี่ยมีแนวโน้มต่ำที่สุด

สุพานี กอหมั่นศิลป์ (2543) ศึกษาความเชื่อมั่นของแบบประเมินทักษะกีฬาบาสเกตบอล ที่ใช้ผู้ประเมินจำนวนต่างกัน ประเมินนักเรียนจำนวนต่างกัน ที่มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าความเชื่อมั่นของแบบประเมินทักษะกีฬาบาสเกตบอลที่ใช้ผู้ประเมินจำนวนต่างกัน ประเมินนักเรียนจำนวนต่างกัน โดยกลุ่มตัวอย่างที่ใช้เป็นนักเรียนชั้นมัธยมศึกษาปีที่ 3 ภาคเรียนที่ 2 ปีการศึกษา 2542 จาก 4 โรงเรียน จำนวน 180 คน ในจังหวัดราชบุรี จากการสุ่มแบบสองขั้นตอน โดยแบ่งเป็น 3 กลุ่ม คือ กลุ่มนักเรียนเข้ารับการประเมินครั้งละ 1 คน จำนวน 30 คน กลุ่มเข้ารับการประเมินครั้งละ 2 คน 60 คน กลุ่มเข้ารับการประเมินครั้งละ 3 คน จำนวน 90 คน ผลการวิจัยพบว่า การเปรียบเทียบค่าความเชื่อมั่นแบบประเมินทักษะกีฬาบาสเกตบอลที่ใช้ผู้ประเมินจำนวนต่างกัน คือ จำนวนหนึ่งคน สองคน และสามคน ประเมินนักเรียนครั้งละ 1 คน 2 คน และ 3 คน ค่าความเชื่อมั่นเฉลี่ยของผู้ประเมินจำนวน 1 คน มีค่าต่ำสุด และ ค่าความเชื่อมั่นเฉลี่ยของผู้ประเมิน จำนวน 3 คน มีค่าสูงสุด เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่น พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .01 และการเปรียบเทียบค่าความเชื่อมั่นของแบบประเมินทักษะกีฬาบาสเกตบอลที่ประเมินนักเรียนจำนวนต่างกัน คือ ครั้งละหนึ่งคน สองคน และสามคน เมื่อใช้ผู้ประเมินจำนวน 1 คน 2 คน และ 3 คน ค่าความเชื่อมั่นเฉลี่ยของนักเรียนที่เข้ารับการประเมินครั้งละ 1 คน มีค่าต่ำสุด และค่าความเชื่อมั่นเฉลี่ยของนักเรียนที่เข้ารับการประเมินครั้งละ 3 คน มีค่าสูงสุด เมื่อทดสอบความแตกต่างของค่าความเชื่อมั่น พบว่า มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

นิศารัตน์ คงสวัสดิ์ (2544) ศึกษาผลการวิเคราะห์ค่าความเชื่อมั่นของผู้ตรวจแบบทดสอบความเรียง ที่มีจำนวนผู้ตรวจ และวิธีการตรวจต่างกัน ที่มีจุดมุ่งหมายสำคัญเพื่อศึกษาและเปรียบเทียบผลการวิเคราะห์ค่าความเชื่อมั่นของผู้ตรวจแบบทดสอบความเรียงที่มีผู้ตรวจจำนวนต่างกัน คือ สองคน สามคน และสี่คน เมื่อใช้การตรวจวิธีเดียวกัน และใช้การตรวจต่างวิธีกัน คือ การตรวจวิธีวิเคราะห์ย่อย (Analytic Method) และ การตรวจวิธีประเมินรวม (Holistic Method) เมื่อมีผู้ตรวจจำนวนเท่ากัน กลุ่มตัวอย่างในการวิจัย เป็นนักเรียนชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาเอกชน จำนวน 60 คน และ ผู้ตรวจแบบทดสอบความเรียงเป็นครูผู้สอนวิชาสร้างเสริมประสบการณ์ชีวิต ชั้นประถมศึกษาปีที่ 5 ของโรงเรียนสังกัดสำนักงานคณะกรรมการการศึกษาเอกชน จำนวนสี่คน ที่มีความสามารถตามเกณฑ์ที่กำหนดตรวจแบบทดสอบความเรียงจำนวนห้าข้อ ด้วยการตรวจวิธีวิเคราะห์ย่อย และวิธีประเมินรวม ทั้งสี่คนหาค่าความเชื่อมั่นของผู้ตรวจโดยวิเคราะห์หาค่าดัชนีความสอดคล้องระหว่างผู้ตรวจ (Rat Agreement Indexes : RAI) ได้ผลการวิจัยดังนี้

1. การเปรียบเทียบความแตกต่างของค่าเฉลี่ยความเชื่อมั่นของผู้ตรวจแบบทดสอบความเรียง ที่มีผู้ตรวจจำนวนต่างกัน คือ สองคน สามคน และสี่คน โดยใช้การตรวจวิธีเดียวกัน พบว่า มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ทั้งการตรวจวิธีวิเคราะห์ย่อย และการตรวจวิธีประเมินรวม

2. การเปรียบเทียบความแตกต่างของค่าเฉลี่ยความเชื่อมั่นของผู้ตรวจแบบทดสอบความเรียง ที่มีผู้ตรวจจำนวนเท่ากัน เมื่อใช้การตรวจต่างวิธีกัน คือ การตรวจวิธีวิเคราะห์ย่อย และการตรวจวิธีประเมินรวม พบว่า มีความแตกต่างกันไม่มีนัยสำคัญ ทั้งสามกรณี คือ เมื่อมีผู้ตรวจจำนวนสองคน สามคน และสี่คน

จากการศึกษางานวิจัยที่เกี่ยวข้องทั้งภายในประเทศและต่างประเทศ จะเห็นได้ว่าจำนวนผู้ประเมินให้คะแนนมีผลให้ความเชื่อมั่นระหว่างการประเมินแตกต่างกันไป ดังนั้นผู้วิจัยจึงสนใจที่จะศึกษาหาความเชื่อมั่นของการประเมินผลงานศิลปะด้วยแบบประเมินเมื่อคำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก สูตร K ของสก็อต และสูตร RAI ของเบอร์รี - สต็อก และคณะ ว่าจะแปรเปลี่ยนไปอย่างไร ถ้ามีเครื่องมือและสูตรการคำนวณค่าความเชื่อมั่นที่ต่างกัน

สมมติฐานของการวิจัย

1. ค่าความเชื่อมั่นที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) มีค่าแตกต่างกัน

2. ค่าความเชื่อมั่นที่คำนวณด้วยสูตร K ของสก็อต (Scott's, 1971) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) มีค่าแตกต่างกัน

3. ค่าความเชื่อมั่นที่คำนวณด้วยสูตร RAI ของเบอร์รี - สต็อก และคณะ (Judith A. Burry - Stock, Cecelia, Brad S. Chissom) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) มีค่าแตกต่างกัน

บทที่ 3

วิธีดำเนินการวิจัย

การกำหนดประชากรและเลือกกลุ่มตัวอย่าง

การศึกษาวิจัยครั้งนี้ มีความมุ่งหมายเพื่อศึกษาความเชื่อมั่นของการประเมินว่าจะมีค่าความเชื่อมั่นที่แตกต่างกันหรือไม่ เมื่อคำนวณจากสูตรที่ต่างกันสามสูตร คือ สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) สูตร K ของสก็อต (Scott's, 1971) และสูตร RAI ของเบอร์รี – สต็อก และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom)

ผู้ประเมินที่ใช้ในการวิจัย

ผู้ประเมิน คือ ครูผู้สอนศิลปศึกษาที่จบปริญญาตรีด้านศิลปศึกษาและมีประสบการณ์การสอนด้านศิลปศึกษาอย่างน้อยห้าปี สังกัดโรงเรียนในสำนักงานเขตคลองสามวา กรุงเทพมหานคร ซึ่งเลือกมาจำนวนห้าคน จากผู้สมัครใจ มาทำการประเมินผลงานศิลปะของนักเรียน จำนวน 30 ภาพ ดังรายละเอียดในตาราง 7

ตาราง 7 จำนวนผู้ประเมินจำแนกตามโรงเรียน

โรงเรียน	จำนวนครูสอนวิชาศิลปศึกษา
1. โรงเรียนประชากรราษฎร์อุปถัมภ์	1
2. โรงเรียนสุเหร่าคลองสี	1
3. โรงเรียนสุเหร่าคลองหนึ่ง	1
4. โรงเรียนวัดสุทธิสะอาด	1
5. โรงเรียนสุเหร่าสามวา	1
รวม	5

เครื่องมือที่ใช้ในการวิจัย

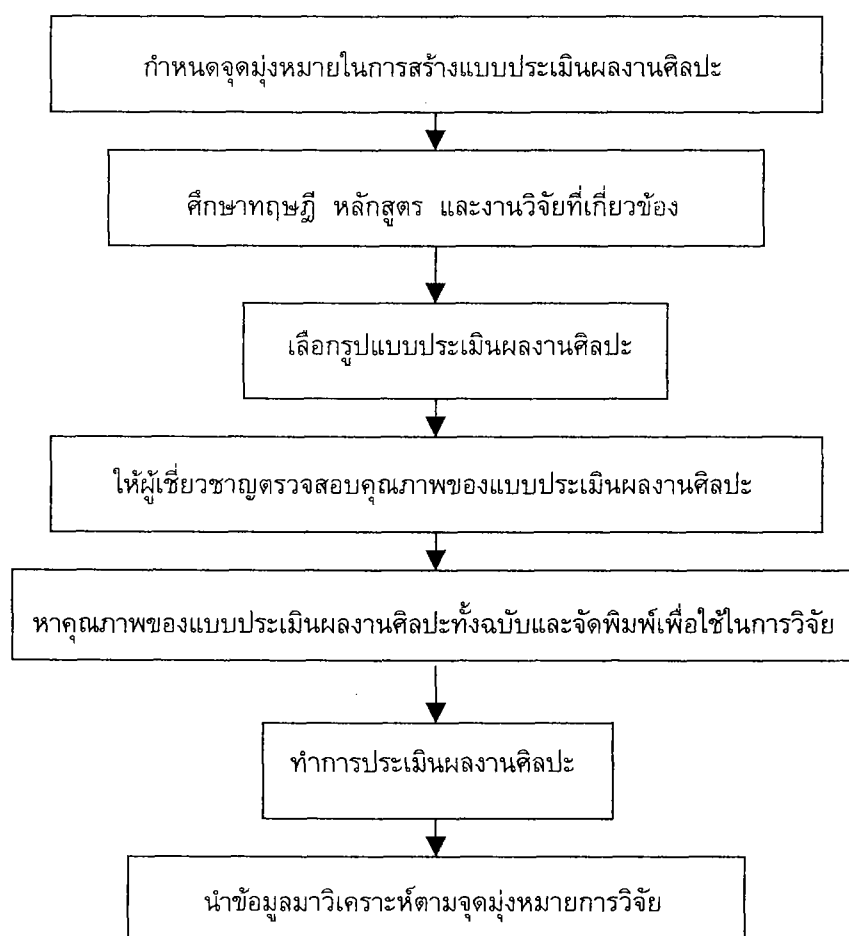
เครื่องมือที่ใช้ในการวิจัยครั้งนี้ เป็นการประเมินผลงานศิลปะที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยกและแบบเกณฑ์รวม ดังรายละเอียดของประเด็นประเมินและเกณฑ์การให้คะแนนด้านต่าง ๆ ดังต่อไปนี้

1. การเน้นจุดเด่นของภาพ หมายถึง ภาพเด่นสะดุดตาและมีการใช้สีสะดุดตาชัดเจน
2. ความสมดุลของภาพ หมายถึง มีการจัดวางภาพเหมาะสมกับหน้ากระดาษและภาพมีระยะใกล้ – ไกลมีการผสมผสานกลมกลืนของรูปและพื้นหลัง
3. การระบายสี หมายถึง มีการระบายสีภาพอยู่ในกรอบสะอาดเรียบร้อย สีเรียบ กลมกลืน

4. การใช้สี หมายถึง ภาพให้แสงและเงาถูกต้องและแยกความสัมพันธ์ภาพ และพื้นหลังชัดเจน ให้สีเหมือนธรรมชาติ
5. ความคิดสร้างสรรค์ หมายถึง ภาพมีแนวคิดแปลกใหม่มีการสื่อความหมายของการใช้ความคิดของตนเอง และไม่ซ้ำภาพของผู้อื่น
6. ความประณีต หมายถึง ภาพมีความละเอียดมีความชัดเจนของภาพและสามารถมองเห็นความงดงามของภาพ
7. ความสะอาด หมายถึง ภาพมีการระบายสีอยู่ในกรอบ ไม่เลอะเทอะ และมีความเป็นระเบียบเรียบร้อย

การสร้างเครื่องมือที่ใช้ในการวิจัย

ผู้วิจัยได้สร้างเครื่องมือวัดผลงานศิลปะเพื่อใช้ในการเก็บรวบรวมข้อมูล จำนวน 2 ฉบับ มีประเด็นประเมิน 7 ประเด็น ตามขั้นตอนการสร้างแบบประเมินผลงานศิลปะ มีรายละเอียดตามลำดับขั้นดังนี้



ภาพประกอบ 1 แสดงลำดับขั้นตอนการสร้างแบบประเมินผลงานศิลปะ

จากภาพประกอบ แสดงขั้นตอนการสร้างแบบประเมินผลงานศิลปะมีรายละเอียดดำเนินการสร้างตามขั้นตอน ดังนี้

1. กำหนดจุดมุ่งหมายในการสร้างแบบประเมินผลงานศิลปะ

1.1 เพื่อสร้างแบบประเมินผลงานศิลปะพร้อมทั้งกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

1.2 เพื่อนำคะแนนจากแบบประเมินไปใช้ในการศึกษาความเชื่อมั่นของการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) ที่มีสูตรการหาความเชื่อมั่นต่างกัน

2. ศึกษาทฤษฎี หลักสูตร และงานวิจัยที่เกี่ยวข้อง

2.1 ศึกษาทฤษฎีและเอกสารงานวิจัยที่เกี่ยวข้องกับความหมาย ลักษณะ ประเภท การสร้างแบบประเมิน และวิธีการประเมินแบบประเมินผลงานศิลปะ

2.2 ศึกษาโครงสร้างหลักสูตร คู่มือการประเมินผลงานศิลปะ งานวิจัยที่เกี่ยวข้องในระดับประถมศึกษา

2.3 ศึกษาวิธีการหาค่าความเชื่อมั่นของการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

3. เลือกรูปแบบการประเมินเป็นแบบค่าคะแนน 4 ระดับ คือ ระดับคะแนน 4 หมายถึง ดีมาก ระดับคะแนน 3 หมายถึง ดี ระดับคะแนน 2 หมายถึง พอใช้ ระดับคะแนน 1 หมายถึง ใช้ไม่ได้

4. ให้ผู้เชี่ยวชาญตรวจสอบคุณภาพของแบบประเมินผลงานศิลปะด้านความเที่ยงตรงเชิงพินิจ (Face Validity) โดยผู้เชี่ยวชาญด้านศิลปะจำนวนสามคนและทางด้านวัดผล จำนวนสองคน รวมทั้งหมดห้าคน เป็นผู้พิจารณาตรวจสอบในด้านต่าง ๆ ต่อไปนี้คือ

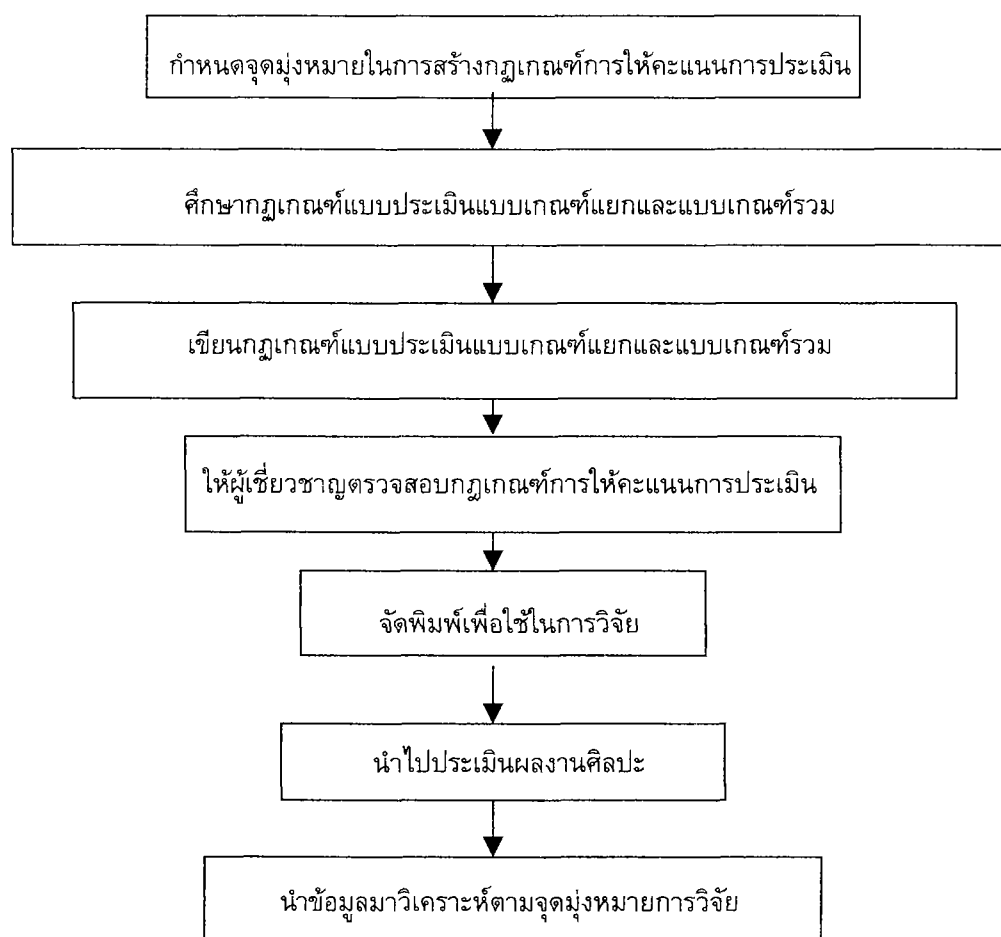
- 1) ความสอดคล้องของข้อประเมินกับเนื้อหาที่ต้องการประเมิน
- 2) การเขียนข้อประเมินได้ชัดเจน ตรงกับสิ่งที่ต้องการประเมิน
- 3) ความยากและความเหมาะสมทางภาษา
- 4) คู่มือการประเมิน ตามวิธีการประเมินให้คะแนนทั้ง 2 แบบ

5. นำผลมาวิเคราะห์หาค่าความเที่ยงตรงโดยใช้ดัชนีความสอดคล้อง (*IOC*) โดยรวมความเห็นในช่อง เห็นด้วย (+1) ไม่แน่ใจ (0) และไม่เห็นด้วย (-1) แล้วหาค่าเฉลี่ยตามวิธีของโรวินลลีและแฮมเบลตัน คัดเลือกข้อที่มีค่า *IOC* ตั้งแต่ 0.5 ขึ้นไป จำนวน 7 ประเด็น ซึ่งมี *IOC* เท่ากับ 1.0 ทั้งหมด ดังแสดงในภาคผนวก

6. นำแบบประเมินผลงานศิลปะไปประเมินภาพศิลปะจำนวน 10 ภาพ โดยให้ประเมินทีละภาพด้วยแบบประเมินแบบเกณฑ์แยกก่อน หลังจากนั้น 1 สัปดาห์ ให้ประเมินด้วยแบบประเมินแบบเกณฑ์รวม และรวบรวมผลการประเมินมาทำการวิเคราะห์หาค่าอำนาจจำแนก (r_{xy}) ของแบบประเมินแบบเกณฑ์แยก และค่าอำนาจจำแนก (r_{xy}) ของแบบประเมินแบบเกณฑ์รวมได้ค่าอำนาจจำแนกตั้งแต่ .900 ถึง 1.000 ดังแสดงในตารางภาคผนวก

7. นำประเด็นการประเมินผลงานศิลปะไปให้ผู้ประเมินทำการประเมินผลงานศิลปะจำนวน 30 ภาพ แล้วนำผลของแบบประเมินมาวิเคราะห์ตามจุดมุ่งหมายของการวิจัยต่อไป

ผู้วิจัยได้สร้างกฎเกณฑ์การให้คะแนนการประเมินทั้งสองวิธี มีรายละเอียดตามลำดับขั้นดังนี้



ภาพประกอบ 2 แสดงลำดับขั้นการสร้างกฎเกณฑ์การให้คะแนนการประเมินผลงานศิลปะ

1. กำหนดจุดมุ่งหมายในการสร้างกฎเกณฑ์การให้คะแนนการประเมิน
ในการวิจัยครั้งนี้ผู้วิจัยได้สร้างกฎเกณฑ์แบบประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)
2. ศึกษากฎเกณฑ์แบบประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)
3. เขียนกฎเกณฑ์แบบประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)
4. นำแบบประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) ให้ผู้เชี่ยวชาญตรวจสอบคุณภาพด้านเนื้อหา ความสอดคล้องและภาษามีความเหมาะสมกับประเด็นที่ประเมินหรือไม่ โดยมีผู้เชี่ยวชาญด้านศิลปศึกษาจำนวนสามคนและทางด้านวัดผลจำนวนสองคน รวมทั้งหมดห้าคน

5. จัดพิมพ์รายละเอียดเพื่อนำไปใช้ในการวิจัย

6. นำกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric)

และแบบเกณฑ์รวม (Holistic Scoring Rubric) ไปประเมินผลงานศิลปะและรวบรวมผลการประเมินมาทำการวิเคราะห์ตามจุดมุ่งหมายของการวิจัยต่อไป

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบประเมินผลงานศิลปะ ซึ่งมีรายละเอียดดังนี้

คำชี้แจง ให้ท่านประเมินผลงานภาพศิลปะตามรายการที่กำหนดให้แล้วให้คะแนนตามความคิดเห็นของท่านตามกฎเกณฑ์ที่กำหนดให้ลงในตารางแต่ละภาพ

กฎเกณฑ์การให้คะแนนการประเมินภาพระบายสีแบบเกณฑ์รวม (Holistic Scoring Rubric)

ระดับคุณภาพ	คำอธิบาย
4	- ภาพเด่นสะดุดตา มีการจัดวางได้เหมาะสมกับกระดาษ แสดงระยะใกล้-ไกล การระบายสีกลมกลืน ให้แสงและเงาถูกต้องเหมือนธรรมชาติ และมีแนวคิดแปลกใหม่เป็นของตัวเอง สามารถมองเห็นได้ชัดเจนงดงามเป็นระเบียบ
3	- ภาพเด่นสะดุดตา จัดวางได้เหมาะสมกับกระดาษ ระบายสีออกนอกเส้นบางแต่อยู่ในกรอบ และ ให้แสงและเงาไม่ถูกต้อง ให้สีใกล้เคียงธรรมชาติ มีความแปลกใหม่ สื่อความหมายของตัวเอง มีความชัดเจนของภาพ
2	- ภาพมีการใช้สีไม่เหมาะสม จัดวางภาพเหมาะสมกับกระดาษ ไม่มีระยะใกล้-ไกล ระบายสีเลอะเทอะ ไม่แสดงแสงและเงา ให้สีผิดธรรมชาติ มีแนวคิดแปลกใหม่ แต่ไม่มีความชัดเจน ไม่มีความเป็นระเบียบของภาพ
1	- ภาพไม่ชัดเจน การวางภาพไม่มีความสมดุล มีการระบายสีกระจาย เลอะเทอะทั่วไป ไม่แสดงแสงและเงา ให้สีผิดธรรมชาติ ไม่มีแนวคิดของตัวเอง ไม่มีความละเอียดของภาพ ไม่มีความสะอาด

แบบประเมินผลงานศิลปะแบบเกณฑ์รวม (Holistic Scoring Rubric)

ประเด็นประเมิน	ภาพที่	ระดับคุณภาพคะแนน			
		4	3	2	1
- การเน้นจุดเด่นของภาพ	000				
- ความสมดุลของภาพ	001				
- การระบายสี	002				
- การใช้สี	003				
- ความคิดสร้างสรรค์	004				
- ความประณีต	005				
- ความสะอาด	006				

กฎเกณฑ์ของการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric)

กฎเกณฑ์การให้คะแนนการประเมินภาพระบายสีแบบเกณฑ์แยก (Analytic Scoring Rubric)

ประเด็นประเมิน	คำอธิบายระดับคุณภาพ			
	4	3	2	1
1. การเน้นจุดเด่นของภาพ	- ภาพเด่นสะดุดตาและมีการใช้สีสะดุดตาชัดเจน	- ภาพเด่นสะดุดตาแต่การใช้สีกระจายจนไม่มีจุดเด่น	- มีการใช้สีมากหรือน้อยไปไม่เหมาะสมกับภาพ	- ภาพไม่ชัดเจนและการให้สีน้อยมาก
2. ความสมดุลของภาพ	- มีการจัดวางภาพเหมาะสมกับหน้ากระดาษ และภาพมีระยะใกล้ – ไกล มีการผสมผสานกลมกลืนของรูปและพื้นหลัง	- มีการจัดวางภาพเหมาะสมกับหน้ากระดาษ และมีการใช้ระยะใกล้ – ไกลของภาพ	- มีการจัดวางภาพเหมาะสมกับหน้ากระดาษแต่ขาดระยะใกล้ – ไกลของภาพ	- ภาพไม่มีลักษณะของการจัดวางภาพที่มีความสมดุลเลย
3. การระบายสี	- มีการระบายสีภาพอยู่ในกรอบสะอาดเรียบร้อย สีเรียบกลมกลืน	- มีการระบายสีภาพออกนอกเส้นเป็นส่วนน้อย สีส่วนใหญ่เรียบ	- มีการระบายสีภาพออกนอกกรอบเลอะเทอะ สีไม่เรียบ	- มีการระบายสีภาพอย่างไม่มีจุดมุ่งหมาย
4. การใช้สี	- ภาพให้แสงและเงาถูกต้อง และแยกความสัมพันธ์ภาพและพื้นหลังชัดเจน ให้สีเหมือนธรรมชาติ	- ภาพให้แสงและเงา แต่ไม่ถูกต้อง และแยกความสัมพันธ์ภาพและพื้นหลังชัดเจน มีการให้สีใกล้เคียงธรรมชาติ	- ภาพไม่แสดงแสงและเงา แต่ในภาพมีสีพื้นหลังกลมกลืนมีการให้สีผิดธรรมชาติ	- ภาพไม่แสดงแสงและเงา มีสีพื้นหลังกลมกลืนไม่เด่นชัด มีการให้สีผิดธรรมชาติ
5. ความคิดสร้างสรรค์	- ภาพมีแนวคิดแปลกใหม่ มีการสื่อความหมายของการใช้ความคิดของตนเองและไม่ซ้ำภาพผู้อื่น	- ภาพมีแนวคิดแปลกใหม่ มีการสื่อความหมายของการใช้ความคิดของตนเอง แต่ซ้ำภาพผู้อื่น	- ภาพมีแนวคิดแปลกใหม่ ไม่มีการสื่อความหมายของการใช้ความคิดของตนเอง	- ภาพไม่มีแนวคิดแปลกใหม่ และซ้ำกับภาพผู้อื่น
6. ความประณีต	- ภาพมีความละเอียดมีความชัดเจนของภาพและสามารถมองเห็นความงามของภาพ	- ภาพมีความละเอียดมีความชัดเจนของภาพแต่ขาดความงดงาม	- ภาพมีความละเอียดแต่ไม่มีความชัดเจนของภาพ	- ไม่มีความละเอียดของภาพและไม่มีภาพชัดเจนของภาพ

กฎเกณฑ์การให้คะแนนการประเมินภาพระบายสีของแบบเกณฑ์แยก (Analytic Scoring Rubrics)
(ต่อ)

ประเด็นประเมิน	คำอธิบายระดับคุณภาพ			
	4	3	2	1
7. ความสะอาด	- ภาพมีการระบายสีอยู่ในกรอบไม่เลอะเทอะและมีความเป็นระเบียบเรียบร้อย	- ภาพมีการระบายสีอยู่ในกรอบไม่เลอะเทอะแต่ขาดความเป็นระเบียบเรียบร้อย	- ภาพมีการระบายสีอยู่ในกรอบ เลอะเทอะและไม่มีความเป็นระเบียบ	- ภาพมีการระบายสีออกนอกกรอบเลอะเทอะ

แบบประเมินผลงานศิลปะของแบบประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric)

คำชี้แจง โปรดใส่เครื่องหมาย ✓ ลงในช่องคะแนนที่ตรงกับการประเมินผลงานศิลปะของท่าน โดยเทียบกับกฎเกณฑ์การให้คะแนนที่กำหนด

ภาพที่ 0	ประเด็นประเมิน	ระดับคุณภาพ			
		คะแนน			
		4	3	2	1
1. การเน้นจุดเด่นของภาพ					
2. ความสมดุลของภาพ					
3. การระบายสี					
4. การใช้สี					
5. ความคิดสร้างสรรค์					
6. ความประณีต					
7. ความสะอาด					
คะแนนรวม					

ภาพที่ 00	ประเด็นประเมิน	ระดับคุณภาพ			
		คะแนน			
		4	3	2	1
1. การเน้นจุดเด่นของภาพ					
2. ความสมดุลของภาพ					
3. การระบายสี					
4. การใช้สี					
5. ความคิดสร้างสรรค์					
6. ความประณีต					
7. ความสะอาด					
คะแนนรวม					

เกณฑ์การให้คะแนน

- คะแนน 1 – 7 คะแนน ระดับคุณภาพ 1
 คะแนน 8 – 14 คะแนน ระดับคุณภาพ 2
 คะแนน 15 – 21 คะแนน ระดับคุณภาพ 3
 คะแนน 22 – 28 คะแนน ระดับคุณภาพ 4

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลในการวิจัย โดยใช้แบบประเมินผลงานศิลปะดังนี้

1. ขอนหนังสือจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อความร่วมมือกับโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง
2. ติดต่อประสานงาน และขอความร่วมมือผู้บริหารโรงเรียนที่เลือกกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลา สถานที่ และวิธีการประเมิน และชี้แจงครูผู้ประเมินที่ร่วมในการวิจัยครั้งนี้ ให้เกิดความเข้าใจเป็นอย่างดี
3. ชี้แจงให้ผู้ประเมินทราบถึงจุดมุ่งหมายของการวิจัยและความสำคัญของการวิจัยนี้ ก่อนดำเนินการประเมินผลงานศิลปะ
4. นำภาพผลงานศิลปะที่ระบายด้วยสีชอล์กจำนวน 30 ภาพ ที่ให้ครูผู้สอนวิชาศิลปะศึกษาทำการคัดเลือกผลงานของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่วาดภาพในเรื่องวันที่ฉันไปเที่ยว จากจำนวน 108 ภาพ เหลือจำนวน 30 ภาพ ให้ผู้ประเมินทำการประเมินตามจุดมุ่งหมายของการวิจัยดังนี้
 - 4.1 จัดภาพผลงานศิลปะจำนวน 30 ภาพ สำหรับการประเมินด้วยผู้ประเมินห้าคนทำการประเมินโดยใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม (Holistic Scoring) ตามคู่มือการประเมินที่ละภาพจนครบทุกภาพ และทุกคน
 - 4.2 หลังจากนั้น 1 สัปดาห์ จัดภาพผลงานศิลปะจำนวน 30 ภาพ สำหรับการประเมินจำนวนห้าคนประเมินมาตรฐานประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) ตามคู่มือการประเมินที่ไว้ให้กับผู้ประเมินทุกคน พร้อมทั้งชี้แจงให้ผู้ประเมินภาพผลงานศิลปะที่ละภาพให้ครบทุกคน และจนครบทุกภาพ
5. นำผลการประเมินมาทำการวิเคราะห์ข้อมูล ดังนี้
 - 5.1 คำนวณหาค่าสถิติพื้นฐาน ($\bar{X}$, $S.D.$ และ $C.V.$)
 - 5.2 สถิติที่ใช้หาคุณภาพของเครื่องมือวัด
 - 5.3 คำนวณหาค่าของสูตร โดยใช้สูตร P_0 ของแฮมเบิลตันและโนวิก สูตร K ของสก็อต และสูตร RAI ของเบอร์รี – สต็อกส์ และคณะ ที่มีรูปแบบการคำนวณตามลำดับ ดังนี้
 - 5.3.1 นำคะแนนผลการประเมินของผู้ประเมินครั้งที่ 1 และครั้งที่ 2 ทั้ง 5 คน ที่ประเมินโดยใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม แล้วคำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก
 - 5.3.2 นำคะแนนผลการประเมินของผู้ประเมินครั้งที่ 1 และครั้งที่ 2 ทั้ง 5 คน ที่ประเมินโดยใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก แล้วคำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก
 - 5.3.3 นำคะแนนผลการประเมินของผู้ประเมินจำนวนสองคน คือ ผู้ประเมินคนที่ 1 กับคนที่ 2, ผู้ประเมินคนที่ 1 กับคนที่ 3, ผู้ประเมินคนที่ 1 กับคนที่ 4, ผู้ประเมินคนที่ 1 กับคนที่ 5, ผู้ประเมินคนที่ 2 กับคนที่ 3, ผู้ประเมินคนที่ 2 กับคนที่ 4, ผู้ประเมินคนที่ 2 กับคนที่ 5, ผู้ประเมินคนที่ 3 กับคนที่ 4, ผู้ประเมินคนที่ 3 กับคนที่ 5, ผู้ประเมินคนที่ 4 กับคนที่ 5 ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม แล้วคำนวณด้วยสูตร K ของสก็อต
 - 5.3.4 นำคะแนนผลการประเมินของผู้ประเมินจำนวนสองคน คือ ผู้ประเมินคนที่ 1 กับคนที่ 2, ผู้ประเมินคนที่ 1 กับคนที่ 3, ผู้ประเมินคนที่ 1 กับคนที่ 4, ผู้ประเมินคนที่ 1 กับคนที่ 5, ผู้ประเมินคนที่ 2 กับคนที่ 3, ผู้ประเมินคนที่ 2 กับคนที่ 4, ผู้ประเมินคนที่ 2 กับคนที่ 5, ผู้ประเมินคนที่ 3 กับคนที่ 4, ผู้ประเมินคนที่ 3 กับคนที่ 5

กับคนที่ 5, ผู้ประเมินคนที่ 4 กับคนที่ 5 ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก แล้วคำนวณด้วยสูตร K ของสก็อต

5.3.5 นำผลการประเมินของผู้ประเมินทั้ง 5 คน ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม แล้วคำนวณด้วยสูตร RAI ของเบอร์รี – สติอจค์ และคณะ

5.3.6 นำผลการประเมินของผู้ประเมินทั้ง 5 คน ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก แล้วคำนวณด้วยสูตร RAI ของเบอร์รี – สติอจค์ และคณะ

5.4 ทดสอบสมมติฐานของการวิจัย

ทดสอบความแตกต่างของสูตรและกฎเกณฑ์ที่ใช้ประเมิน ตามสมมติฐานของการวิจัยโดยใช้สูตร UX_1 ของวูดรอฟฟ์ และเฟลด์ท์ (Woodruff and feldt. 1986 : 393 – 413) อีกครั้ง

การจัดกระทำและการวิเคราะห์ข้อมูล

สถิติที่ใช้ แบ่งออกเป็น 3 ลักษณะ คือ

1. หาค่าสถิติพื้นฐาน

1.1 หาค่าเฉลี่ย (Mean)

1.2 หาค่าเบี่ยงเบนมาตรฐาน (Standard deviation)

1.3 หาค่าสัมประสิทธิ์ความแปรผัน (Coefficient of Variation)

2. สถิติที่ใช้หาคุณภาพของเครื่องมือวัด

2.1 หาค่าอำนาจจำแนกของแบบประเมิน โดยใช้สูตรสหสัมพันธ์แบบเพียร์สัน (Pearson Product Moment Correlation : r_{xy}) (บุญเชิด ภิญโญนนตพงษ์. 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 ดัชนีอำนาจจำแนกของการประเมิน

N แทน จำนวนภาพ

$\sum X$ แทน ผลรวมของคะแนนของภาพที่ประเมิน

$\sum Y$ แทน ผลรวมของคะแนนเฉลี่ยของภาพประเมิน

$\sum XY$ แทน ผลรวมของผลคูณของคะแนน X และ Y

$\sum X^2$ แทน ผลรวมกำลังสองของคะแนน X

$\sum Y^2$ แทน ผลรวมกำลังสองของคะแนน Y

2.2 หาดัชนีความสอดคล้องของผู้เชี่ยวชาญ (Index of Consistency : IOC)
(บุญเชิด ภิญโญนนตพงษ์. 2527 : 69) ดังนี้

$$IOC = \frac{\sum R}{N}$$

เมื่อ IOC แทน ดัชนีความสอดคล้องมีค่าอยู่ระหว่าง -1 ถึง +1
 $\frac{\sum R}{N}$ แทน ผลรวมของการพิจารณาของผู้เชี่ยวชาญ
 N แทน จำนวนผู้เชี่ยวชาญ

3. สถิติที่ใช้ทดสอบสมมติฐาน

3.1 สูตร P_0 ของแฮมเบิลตันและโนวิก (บุญเชิด ภิญโญนนตพงษ์. 2527 ; Subkoviak. 1976 : 266 citing Hambleton and Novick. 1973)

$$p_0 = \sum_{i=1}^2 P_{ii} = P_{11} + P_{22}$$

P_0 แทน สัดส่วนของความสอดคล้องในการประเมินเพื่อจำแนก
 P_{11} แทน สัดส่วนของผู้ประเมินว่าประเมินตรงกันทั้งสองครั้ง
 P_{22} แทน สัดส่วนของผู้ประเมินว่าประเมินไม่ตรงกันทั้งสองครั้ง

3.2 สูตร K ของสก็อต (บุญเชิด ภิญโญนนตพงษ์. 2545 ; Ober, Richard L., Bentley, Ernest I., and Miller, Edith. 1971 : 80)

$$\text{ความเชื่อมั่น (Reliability) (} K \text{)} = \frac{p_0 - p_e}{1.00 - p_e}$$

P_0 แทน ความสอดคล้องระหว่างผู้ประเมิน 2 คน
 P_e แทน ความสอดคล้องระหว่างผู้ประเมิน 2 คนที่เกิดขึ้นในตัวอย่างโดย
 ความบังเอิญ

3.3 สูตร RAI ของเบอร์รี่ – สต็อกและคณะ (Burry – Stock and others. 1996)
 แบบที่ 4 คือ กรณีที่ประเมินพฤติกรรมเดียว ผู้ประเมิน M คน และกลุ่มตัวอย่างเพียงคนเดียว สูตรที่ใช้
 คำนวณ คือ

$$RAI = 1 - \frac{\sum^M |R_m - \bar{R}|}{(M-1)(I-1)}$$

เมื่อ	R	แทน ผลการประเมินของผู้ประเมินให้คะแนน
	M	แทน จำนวนผู้ประเมินให้คะแนนทั้งหมด
	I	แทน จำนวนช่วงคะแนน (1,2,3,4)

3.4 สถิติทดสอบนัยสำคัญของความแตกต่างของค่าความเชื่อมั่นหลายค่า ซึ่งประมาณจากกลุ่มตัวอย่างกลุ่มเดียว โดยใช้สูตร UX_1 (บุญเชิด ภิญโญนนตพงษ์. 2537 : 57 อ้างอิงมาจาก Woodruff & Feldt. 1986 : 393 – 413)

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

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

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

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

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

$$N_c = N \left[\frac{\bar{n}_h - 1}{\bar{n}_h + 1} \right]$$

$$\bar{n}_h = \frac{M}{\sum (1/n_i)}$$

เมื่อ	UX_1	แทน สถิติทดสอบที่มีการแจกแจงคล้ายกับ χ^2 ที่ $df = m - 1$
	N	แทน จำนวนผู้ประเมิน
	r	แทน ค่าความเชื่อมั่นที่คำนวณได้
	r_{ij}	แทน ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร
	S_u^2	แทน ความแปรปรวนของ U_i
	C_u	แทน ความแปรปรวนร่วมของ U_i
	$\bar{U}$	แทน ค่าเฉลี่ยของคะแนน U_i

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูลและการแปลความหมายของการวิเคราะห์ข้อมูล ในการวิจัยครั้งนี้ เพื่อให้เกิดความเข้าใจตรงกัน ผู้วิจัยจึงได้กำหนดสัญลักษณ์และอักษรย่อที่ใช้ในการวิเคราะห์ข้อมูลดังต่อไปนี้

สัญลักษณ์และอักษรย่อในการวิเคราะห์ข้อมูล

$\bar{X}$	แทน	ค่าเฉลี่ย
$S.D.$	แทน	ค่าเบี่ยงเบนมาตรฐานของคะแนน
$C.V.$	แทน	ค่าสัมประสิทธิ์ความแปรผันของคะแนน
P_0	แทน	ดัชนีความสอดคล้องระหว่างผู้ประเมินของแฮมเบิลตันและโนวิก (Hambleton and Novick. 1973)
K	แทน	ดัชนีความสอดคล้องระหว่างผู้ประเมินของสก็อต (Scott. 1971)
RAI	แทน	ดัชนีความสอดคล้องระหว่างผู้ประเมินของเบอร์รี – สต็อก และคณะ (Burry – Stock and others. 1996)
UX_1	แทน	สถิติทดสอบนัยสำคัญที่มีการแจกแจงแบบไคกำลังสอง (χ^2)
p	แทน	ระดับนัยสำคัญ

การเสนอผลการวิเคราะห์ข้อมูล

การศึกษาค้นคว้าครั้งนี้ ผู้วิจัยได้เสนอผลการวิเคราะห์ข้อมูลตามลำดับ ดังนี้

1. ค่าสถิติพื้นฐานของการประเมินผลงานศิลปะ
2. ค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่คำนวณด้วยสูตร 3 สูตร
3. การเปรียบเทียบค่าความเชื่อมั่นของการประเมินที่มีสูตรต่างกัน
4. การเปรียบเทียบค่าความเชื่อมั่นของการประเมินที่มีรูปแบบประเมินต่างกัน

ผลการวิเคราะห์ข้อมูล

1. ค่าสถิติพื้นฐานของคะแนนการประเมิน

การวิเคราะห์ข้อมูลในขั้นตอนนี้ ผู้วิจัยนำคะแนนการประเมินผลงานศิลปะของผู้ประเมินห้าคนที่ให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) มาวิเคราะห์หาค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์การแปรผันของผู้ประเมินแต่ละคน ดังตาราง 8

ตาราง 8 ค่าสถิติพื้นฐานของคะแนนการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนแบบเกณฑ์แยก และแบบเกณฑ์รวมของผู้ประเมิน

การประเมิน	ผู้ประเมิน คนที่	คะแนนเต็ม	$\bar{X}$	<i>S.D.</i>	<i>C.V.</i> (%)
กฎเกณฑ์การให้คะแนนการ ประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric)	1	4	3.000	0.186	6.187
	2	4	3.467	0.507	14.629
	3	4	3.100	0.421	13.565
	4	4	3.233	0.430	13.306
	5	4	3.000	0.587	19.573
รวม		4	3.160	0.426	13.452
กฎเกณฑ์การให้คะแนนการ ประเมินแบบเกณฑ์รวม (Holistic Scoring Rubric)	1	4	2.700	0.904	33.474
	2	4	2.333	0.607	25.997
	3	4	2.767	0.679	24.536
	4	4	3.300	0.466	14.124
	5	4	2.800	0.610	21.796
รวม		4	2.780	0.653	23.985

ผลการวิเคราะห์ตามตาราง 8 พบว่าการประเมินผลงานศิลปะที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) ทั้งฉบับมีค่าเฉลี่ยเท่ากับ 3.160 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.426 และค่าสัมประสิทธิ์ความแปรผัน เท่ากับ 13.452 %

ส่วนการประเมินผลงานศิลปะที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม (Holistic Scoring Rubric) ทั้งฉบับมีค่าเฉลี่ยเท่ากับ 2.780 คะแนน ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.653 และค่าสัมประสิทธิ์ความแปรผัน เท่ากับ 23.985 % ตามลำดับ

2. ค่าความเชื่อมั่นของการให้คะแนนการประเมินผลงานศิลปะ

การวิเคราะห์ตอนนี้นำผู้วิจัยนำคะแนนการประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยกกับแบบเกณฑ์รวม มาวิเคราะห์หาค่าความเชื่อมั่นโดยใช้สูตร P_0 ของแอมเบลตันและโนวิก สูตร K ของสก็อตและสูตร RAI ของเบอร์รี่ – สติ๊กและคณะ ดังตาราง 9

ตาราง 9 ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะ ที่คำนวณด้วยสูตรต่างกัน

การประเมิน	สูตรการประเมิน		
	P_0	K	RAI
กฎเกณฑ์การให้คะแนนแบบเกณฑ์แยก (Analytic Scoring Rubric)	0.906	0.864	0.873
กฎเกณฑ์การให้คะแนนแบบเกณฑ์รวม (Holistic Scoring Rubric)	0.700	0.769	0.817

ผลการวิเคราะห์ตามตาราง 9 พบว่า การประเมินผลงานศิลปะ เมื่อใช้กฎเกณฑ์การให้คะแนนแบบเกณฑ์แยก สูตร P_0 ของแอมเบลตันและโนวิก ให้ค่าความเชื่อมั่นสูงเป็นอันดับแรก ($P_0 = 0.906$) รองลงมาคือ สูตร RAI ของเบอร์รี – สต็อกและคณะ ($RAI = 0.873$) และสูตร K ของสก็อต ($K = 0.864$) ตามลำดับ

ส่วน เมื่อใช้กฎเกณฑ์การให้คะแนนแบบเกณฑ์รวม สูตร RAI ของเบอร์รี – สต็อกและคณะ ให้ค่าความเชื่อมั่นสูงเป็นอันดับแรก ($RAI = 0.817$) รองลงมาคือ สูตร K ของสก็อต ($K = 0.769$) และสูตร P_0 ของแอมเบลตันและโนวิก ($P_0 = 0.700$) ตามลำดับ

3. การเปรียบเทียบค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่คำนวณด้วยสูตรต่างกัน

การวิเคราะห์ตอนนี้ ผู้วิจัยนำค่าความเชื่อมั่นที่คำนวณด้วยสูตรสามสูตร เมื่อใช้การประเมินรูปแบบเดียวกัน มาเปรียบเทียบกันและทดสอบความแตกต่างโดยใช้สูตร UX_1 ดังตาราง 10

ตาราง 10 การเปรียบเทียบค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่มีสูตรต่างกัน

การประเมิน	สูตรการประเมิน			UX_1	p
	P_0	K	RAI		
กฎเกณฑ์การให้คะแนนแบบเกณฑ์แยก (Analytic Scoring Rubric)	0.906	0.864	0.873	1.075	0.586
กฎเกณฑ์การให้คะแนนแบบเกณฑ์รวม (Holistic Scoring Rubric)	0.700	0.769	0.817	1.637	0.444

ผลการวิเคราะห์ตามตาราง 10 พบว่า สูตร P_0 , K และ RAI ที่คำนวณจากคะแนนการประเมินที่ใช้กฎเกณฑ์การให้คะแนนแบบเกณฑ์แยก ให้ค่าความเชื่อมั่นไม่แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

และที่คำนวณจากคะแนนการประเมินที่ใช้กฎเกณฑ์การให้คะแนนแบบเกณฑ์รวมให้ค่าความเชื่อมั่นไม่แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เช่นเดียวกัน

4 การเปรียบเทียบความเชื่อมั่นของคะแนนการประเมินที่ใช้กฎเกณฑ์การให้คะแนนการประเมินต่างกัน การวิเคราะห์ตอนนี้ ผู้วิจัยนำค่าความเชื่อมั่นของการประเมินที่ใช้กฎเกณฑ์การให้คะแนนต่าง ๆ มาเปรียบเทียบกัน และทดสอบความแตกต่างโดยใช้สูตร UX_1 ดังตาราง 11

ตาราง 11 การเปรียบเทียบค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่มีรูปแบบประเมินต่างกัน

สูตรการประเมิน	รูปแบบการประเมิน		UX_1	p
	กฎเกณฑ์การให้คะแนนแบบ เกณฑ์แยก (Analytic Scoring Rubric)	กฎเกณฑ์การให้คะแนนแบบ เกณฑ์รวม (Holistic Scoring Rubric)		
P_0	0.906	0.700	8.577**	0.004
K	0.864	0.769	1.875	0.167
RAI	0.873	0.817	0.897	0.347

ผลการวิเคราะห์ตามตาราง 11 พบว่า การประเมินผลงานศิลปะที่ใช้กฎเกณฑ์การให้คะแนนแบบเกณฑ์แยกให้ค่าความเชื่อมั่นแตกต่างอย่างมีนัยสำคัญทางสถิติที่ .01 จากการประเมินการให้คะแนนแบบเกณฑ์รวม เมื่อคำนวณด้วยสูตร P_0 แต่เมื่อคำนวณด้วยสูตร K กับ RAI ให้ค่าความเชื่อมั่นไม่ต่างกัน

บทที่ 5

สรุปผล อภิปราย และข้อเสนอแนะ

ความมุ่งหมายของการวิจัย

การศึกษาวิจัยครั้งนี้ มีจุดมุ่งหมาย เพื่อหาค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่มีกฎเกณฑ์ของแบบประเมินต่างกัน จะมีค่าความเชื่อมั่นที่ต่างกันหรือไม่ เมื่อคำนวณจากสูตรที่ต่างกัน โดยมีจุดมุ่งหมายเฉพาะดังนี้

1. เพื่อหาค่าความเชื่อมั่นของการประเมินผลงานศิลปะ เมื่อคำนวณด้วยสูตรและกฎเกณฑ์การให้คะแนนการประเมินต่างกันดังนี้

1.1 ค่าความเชื่อมั่นที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

1.2 ค่าความเชื่อมั่นที่คำนวณด้วยสูตร K ของสก็อต (Scott's, 1971) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

1.3 ค่าความเชื่อมั่นที่คำนวณด้วยสูตร RAI ของเบอร์รี - สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

2. เพื่อเปรียบเทียบค่าความเชื่อมั่นของการประเมินผลงานศิลปะที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) สูตร K ของสก็อต (Scott's, 1971) และสูตร RAI ของเบอร์รี – สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric)

สมมติฐานของการวิจัย

1. ค่าความเชื่อมั่นที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick, 1973) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) มีค่าแตกต่างกัน

2. ค่าความเชื่อมั่นที่คำนวณด้วยสูตร K ของสก็อต (Scott's, 1971) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) มีค่าแตกต่างกัน

3. ค่าความเชื่อมั่นที่คำนวณด้วยสูตร RAI ของเบอร์รี - สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ที่มีกฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) และแบบเกณฑ์รวม (Holistic Scoring Rubric) มีค่าแตกต่างกัน

ผู้ประเมินที่ใช้ในการวิจัย

ผู้ประเมิน คือ ครูผู้สอนศิลปศึกษาที่จบปริญญาตรีด้านศิลปศึกษาและมีประสบการณ์การสอนด้านศิลปศึกษาอย่างน้อย 5 ปี สังกัดโรงเรียนในสำนักงานเขตคลองสามวา กรุงเทพมหานคร ซึ่งเลือกมาจำนวน 5 คน จากผู้สมัครใจ มาทำการประเมินผลงานศิลปะของนักเรียน จำนวน 30 ภาพ

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นมาตรวัดประเมินค่าของการประเมินผลงานศิลปะที่มีการให้คะแนนแบบแยกองค์ประกอบและแบบให้คะแนนภาพรวม ดังรายละเอียดของหัวข้อและเกณฑ์การให้คะแนนด้านต่าง ๆ ดังต่อไปนี้

1. การเน้นจุดเด่นของภาพ หมายถึง ภาพเด่นสะดุดตาและมีการใช้สีสะดุดตาชัดเจน
2. ความสมดุลของภาพ หมายถึง มีการจัดวางภาพเหมาะสมกับหน้ากระดาษและภาพมีระยะใกล้ – ไกลมีการผสมผสานกลมกลืนของรูปและพื้นหลัง
3. การระบายสี หมายถึง มีการระบายสีภาพอยู่ในกรอบสะอาดเรียบร้อย สีเรียบ กลมกลืน
4. การใช้สี หมายถึง ภาพให้แสงและเงาถูกต้องและแยกความสัมพันธ์ภาพ และพื้นหลังชัดเจน ให้สีเหมือนธรรมชาติ
5. ความคิดสร้างสรรค์ หมายถึง ภาพมีแนวคิดแปลกใหม่มีการสื่อความหมายของการใช้ความคิดของตนเอง และไม่ซ้ำภาพของผู้อื่น
6. ความประณีต หมายถึง ภาพมีความละเอียดมีความชัดเจนของภาพและสามารถมองเห็นความงดงามของภาพ
7. ความสะอาด หมายถึง ภาพมีการระบายสีอยู่ในกรอบ ไม่เลอะเทอะ และมีความเป็นระเบียบเรียบร้อย

วิธีดำเนินการเก็บรวบรวมข้อมูล

ผู้วิจัยได้ดำเนินการเก็บรวบรวมข้อมูลในการวิจัย โดยใช้แบบประเมินผลงานศิลปะดังนี้

1. ขออนุญาตจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อความร่วมมือกับโรงเรียนที่ใช้เป็นกลุ่มตัวอย่าง
2. ติดต่อประสานงาน และขอความร่วมมือผู้บริหารโรงเรียนที่เลือกกลุ่มตัวอย่าง เพื่อกำหนดวัน เวลา สถานที่ และวิธีการประเมิน และชี้แจงครูผู้ประเมินที่ร่วมในการวิจัยครั้งนี้ ให้เกิดความเข้าใจเป็นอย่างดี
3. ชี้แจงให้ผู้ประเมินทราบถึงจุดมุ่งหมายของการวิจัยและความสำคัญของการวิจัยนี้ ก่อนดำเนินการประเมินผลงานศิลปะ
4. นำภาพผลงานศิลปะที่ระบายด้วยสีชอล์กจำนวน 30 ภาพ ที่ให้ครูผู้สอนวิชาศิลปศึกษาทำการคัดเลือกผลงานของนักเรียนชั้นประถมศึกษาปีที่ 6 ที่วาดภาพในเรื่องวันที่ฉันไปเที่ยว จากจำนวน 108 ภาพ เหลือจำนวน 30 ภาพ ให้ผู้ประเมินทำการประเมินตามจุดมุ่งหมายของการวิจัยดังนี้

4.1 จัดภาพผลงานศิลปะจำนวน 30 ภาพ สำหรับการประเมินด้วยผู้ประเมินห้าคนทำการประเมินโดยใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม (Holistic Scoring) ตามคู่มือการประเมินที่ละภาพจนครบทุกภาพ และทุกคน

4.2 หลังจากนั้น 1 สัปดาห์ จัดภาพผลงานศิลปะจำนวน 30 ภาพ สำหรับการประเมินจำนวนห้าคนประเมินมาตรฐานประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric) ตามคู่มือการประเมินที่ไว้ให้กับผู้ประเมินทุกคน พร้อมทั้งชี้แจงให้ผู้ประเมินภาพผลงานศิลปะที่ละภาพให้ครบทุกคน และจนครบทุกภาพ

5. นำผลการประเมินมาทำการวิเคราะห์ข้อมูล ดังนี้

5.1 คำนวณหาค่าสถิติพื้นฐาน ($\bar{X}$, $S.D.$ และ $C.V.$)

5.2 สถิติที่ใช้หาคุณภาพของเครื่องมือวัด

5.3 คำนวณหาค่าของสูตร โดยใช้สูตร P_0 ของแฮมเบลตันและโนวิก สูตร K ของสก็อต และสูตร RAI ของเบอร์รี – สต็อกค์ และคณะ ที่มีรูปแบบการคำนวณตามลำดับ ดังนี้

5.3.1 นำคะแนนผลการประเมินของผู้ประเมินครั้งที่ 1 และครั้งที่ 2 ทั้ง 5 คน ที่ประเมินโดยใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม แล้วคำนวณด้วยสูตร P_0 ของแฮมเบลตันและโนวิก

5.3.2 นำคะแนนผลการประเมินของผู้ประเมินครั้งที่ 1 และครั้งที่ 2 ทั้ง 5 คน ที่ประเมินโดยใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก แล้วคำนวณด้วยสูตร P_0 ของแฮมเบลตันและโนวิก

5.3.3 นำคะแนนผลการประเมินของผู้ประเมินจำนวนสองคน คือ ผู้ประเมินคนที่ 1 กับคนที่ 2, ผู้ประเมินคนที่ 1 กับคนที่ 3, ผู้ประเมินคนที่ 1 กับคนที่ 4, ผู้ประเมินคนที่ 1 กับคนที่ 5, ผู้ประเมินคนที่ 2 กับคนที่ 3, ผู้ประเมินคนที่ 2 กับคนที่ 4, ผู้ประเมินคนที่ 2 กับคนที่ 5, ผู้ประเมินคนที่ 3 กับคนที่ 4, ผู้ประเมินคนที่ 3 กับคนที่ 5, ผู้ประเมินคนที่ 4 กับคนที่ 5 ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม แล้วคำนวณด้วยสูตร K ของสก็อต

5.3.4 นำคะแนนผลการประเมินของผู้ประเมินจำนวนสองคน คือ ผู้ประเมินคนที่ 1 กับคนที่ 2, ผู้ประเมินคนที่ 1 กับคนที่ 3, ผู้ประเมินคนที่ 1 กับคนที่ 4, ผู้ประเมินคนที่ 1 กับคนที่ 5, ผู้ประเมินคนที่ 2 กับคนที่ 3, ผู้ประเมินคนที่ 2 กับคนที่ 4, ผู้ประเมินคนที่ 2 กับคนที่ 5, ผู้ประเมินคนที่ 3 กับคนที่ 4, ผู้ประเมินคนที่ 3 กับคนที่ 5, ผู้ประเมินคนที่ 4 กับคนที่ 5 ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก แล้วคำนวณด้วยสูตร K ของสก็อต

5.3.5 นำผลการประเมินของผู้ประเมินทั้ง 5 คน ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์รวม แล้วคำนวณด้วยสูตร RAI ของเบอร์รี – สต็อกค์ และคณะ

5.3.6 นำผลการประเมินของผู้ประเมินทั้ง 5 คน ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยก แล้วคำนวณด้วยสูตร RAI ของเบอร์รี – สต็อกค์ และคณะ

5.4 ทดสอบสมมติฐานของการวิจัย

ทดสอบความแตกต่างของสูตรและกฎเกณฑ์ที่ใช้ประเมิน ตามสมมติฐานของการวิจัยโดยใช้สูตร UX_1 ของวูดรอฟฟ์ และเฟลด์ท์ (Woodruff and Feldt. 1986 : 393 – 413) อีกครั้ง

สรุปผลการวิเคราะห์ข้อมูล

1. ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยกและแบบเกณฑ์รวม มีค่าเท่ากับ 0.906 และ 0.700 ตามลำดับ
2. ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่คำนวณด้วยสูตร K ของสก็อต (Scott) ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยกและแบบเกณฑ์รวม มีค่าเท่ากับ 0.864 และ 0.769 ตามลำดับ
3. ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่คำนวณด้วยสูตร RAI ของเบอร์รี – สต็อก และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ที่ใช้กฎเกณฑ์การให้คะแนนการประเมินแบบเกณฑ์แยกและแบบเกณฑ์รวม มีค่าเท่ากับ 0.873 และ 0.817 ตามลำดับ
4. เปรียบเทียบค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะแต่ละรูปแบบที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) สูตร K ของสก็อต (Scott) และสูตร RAI ของเบอร์รี – สต็อก และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) พบว่า ทั้งสองรูปแบบเมื่อคำนวณด้วยสูตรต่างกัน มีค่าความเชื่อมั่นแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

อภิปรายผล

การศึกษาค้นคว้าครั้งนี้ มีจุดมุ่งหมาย เพื่อหาค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่มีกฎเกณฑ์ของการประเมินที่ต่างกัน จะมีค่าความเชื่อมั่นที่ต่างกันหรือไม่ เมื่อคำนวณจากสูตรที่ต่างกัน ซึ่งผลการศึกษาสามารถนำมาอภิปรายได้ดังนี้

1. ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่คำนวณด้วยสูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) ที่มีรูปแบบของการประเมินต่างกัน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เป็นไปตามสมมติฐาน โดยการให้คะแนนการประเมินแบบเกณฑ์แยกมีค่าความเชื่อมั่นสูงสุด เป็นเพราะประเด็นที่ใช้ในการประเมินผลงานศิลปะมีการเขียนระดับคุณภาพได้ชัดเจน ทำให้ผู้ประเมินสามารถอ่านระดับคุณภาพของแต่ละประเด็นได้เข้าใจมากขึ้นกว่าการให้คะแนนการประเมินแบบเกณฑ์รวม ดังนั้นผลการศึกษาวิจัยครั้งพบว่าการให้คะแนนการประเมินแบบเกณฑ์แยกเป็นรูปแบบที่มีความเหมาะสมในการประเมินผลงานศิลปะที่วาดภาพโดยใช้สีชอล์กมากที่สุด
2. ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่คำนวณด้วยสูตร K ของสก็อต (Scott) ที่มีรูปแบบของการประเมินต่างกัน แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เป็นไปตามสมมติฐาน โดยการให้คะแนนการประเมินแบบเกณฑ์แยกมีค่าความเชื่อมั่นสูงสุด เป็นเพราะประเด็นที่ใช้ในการประเมินผลงานศิลปะมีการเขียนระดับคุณภาพได้ชัดเจน ทำให้ผู้ประเมินสามารถอ่านระดับคุณภาพของแต่ละประเด็นได้เข้าใจมากขึ้นกว่าการให้คะแนนการประเมินแบบเกณฑ์รวม ดังนั้นผลการศึกษาวิจัยครั้งพบว่าการให้คะแนนการประเมินแบบเกณฑ์แยกเป็นรูปแบบที่มีความเหมาะสมในการประเมินผลงานศิลปะที่วาดภาพโดยใช้สีชอล์กมากที่สุด
3. ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะที่คำนวณด้วยสูตร RAI ของเบอร์รี – สต็อก และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ที่มีรูปแบบของการประเมินต่างกัน

แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ เป็นไปตามสมมติฐาน โดยการให้คะแนนการประเมินแบบเกณฑ์แยกมีค่าความเชื่อมั่นสูงสุด เป็นเพราะประเด็นที่ใช้ในการประเมินผลงานศิลปะมีการเขียนระดับคุณภาพได้ชัดเจน ทำให้ผู้ประเมินสามารถอ่านระดับคุณภาพของแต่ละประเด็นได้เข้าใจมากขึ้นกว่าการให้คะแนนการประเมินแบบเกณฑ์รวม ดังนั้นผลการศึกษาวิจัยครั้งพบว่า การให้คะแนนการประเมินแบบเกณฑ์แยกเป็นรูปแบบที่มีความเหมาะสมในการประเมินผลงานศิลปะที่วาดภาพโดยใช้สีชอล์กมากที่สุด

4. ค่าความเชื่อมั่นของแบบประเมินผลงานศิลปะแต่ละรูปแบบ เมื่อคำนวณด้วยสูตรต่างกัน สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) สูตร K ของสก็อต (Scott) และสูตร RAI ของเบอร์รี – สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ ซึ่งเป็นไปตามสมมติฐาน สอดคล้องกับงานวิจัยของเกศเกล้า หมุ่มมิ่ง (2539) ซึ่งพบว่า การประมาณค่าความเชื่อมั่นของมาตรวัดประเมินค่า โดยใช้สูตรต่างกัน จะได้ค่าความเชื่อมั่นที่แตกต่างกัน สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) ที่มีการให้คะแนนการประเมินแบบเกณฑ์แยกนั้น จัดว่าเป็นสูตรที่มีความเหมาะสมของค่าความเชื่อมั่นมากกว่าสูตร K ของสก็อต (Scott) และสูตร RAI ของเบอร์รี – สต็อกค์ และคณะ (Judith A. Burry – Stock, Cecelia, Brad S. Chissom) ซึ่งมีค่าความเชื่อมั่นที่น้อยกว่า ซึ่งผลการศึกษานี้สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) ที่มีการให้คะแนนการประเมินแบบเกณฑ์แยก ให้ค่าความเชื่อมั่นสูง

ข้อเสนอแนะ

ผลการวิจัยครั้งนี้ สามารถสรุปเป็นข้อเสนอในการนำไปใช้และในการวิจัยได้ดังนี้

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

ควรใช้แบบประเมินผลงานศิลปะการให้คะแนนการประเมินแบบเกณฑ์แยก และใช้สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) ในการคำนวณหาค่าความเชื่อมั่น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

ควรศึกษาลักษณะเดียวกันนี้ โดยการจัดผลงานศิลปะเป็นกลุ่มแล้วคำนวณด้วยใช้สูตร P_0 ของแฮมเบิลตันและโนวิก (Hambleton and Novick) ในการคำนวณหาค่าความเชื่อมั่น

บรรณานุกรม

บรรณานุกรม

- กรมวิชาการ. (2539). *การประเมินผลจากสภาพจริง*. กรุงเทพฯ : กรมฯ.
- กองวิชาการ. (2540). *ศิลปศึกษา*. กรุงเทพฯ : สำนักงานคณะกรรมการการประถมศึกษาแห่งชาติ
กระทรวงศึกษาธิการ.
- เกศเกล้า หมูมิ่ง. (2538). *การประมาณค่าความเชื่อมั่นของมาตรวัดแบบประเมินค่า*. ปรินญานิพนธ์ กศม.
(วัดผลการศึกษา) กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- เกษร ธิตะจारी. (2530). *โครงการตำราและเอกสารทางวิชาการคณะครุศาสตร์*. จุฬาลงกรณ์มหาวิทยาลัย.
กรุงเทพฯ : อัดสำเนา
- โครงการพัฒนาทรัพยากรมนุษย์ (ม.ป.ป.). *การประเมินผลโดยใช้แฟ้มสะสมงานดีเด่น*. กระบี่ : โครงการพัฒนาทรัพยากร
มนุษย์ กระทรวงศึกษาธิการ.
- จินตนา ธนวิบูลย์. "การวัดการปฏิบัติ สภาพปัญหาและแนวทางพัฒนา," วารสารการวัดผลการศึกษา. 14 (41) :
43-57 ; กันยายน-ธันวาคม 2535.
- จิราพร อัครสมพงศ์. (2536). *การเปรียบเทียบความสอดคล้องของคะแนนแบบสอบเรียงจากผู้ตรวจและวิธีการตรวจ
ต่างกัน*. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ชวาล แพรัตนกุล. (2508). *เทคนิคการวัดผล*. กรุงเทพฯ : วัฒนาพานิช.
- ชาญวิทย์ จรัสสุทธิอิสร. (2545). *การพัฒนากฎเกณฑ์การให้คะแนนทักษะกระบวนการทางวิทยาศาสตร์ชั้น
บูรณาการ*. ปรินญานิพนธ์กศม.(การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัยมหาวิทยาลัย
ศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- เชาวนา ขวลิตธารง. "การตรวจให้คะแนนการทดสอบภาคปฏิบัติ," การวัดผลการปฏิบัติงาน มหาวิทยาลัย
ศรีนครินทรวิโรฒ ปทุมวัน, 2534.
- ชัยฤทธิ์ ศิลาดเดช. (2540). *การพัฒนาแฟ้มสะสมงานในการประเมินผลการเรียนรู้ภาษาอังกฤษระดับชั้นมัธยมศึกษาปีที่สาม*.
ปรินญานิพนธ์ กศด. (การทดสอบและการวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย. มหาวิทยาลัย
ศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ดำรง วงศ์สว่าง. (2528). *ผลของการทดสอบย่อยโดยใช้แบบทดสอบเลือกตอบ เดิมคำ และเลือกตอบ ที่มี
ต่อความสามารถในการเขียนภาษาไทย ของนักเรียนชั้นมัธยมศึกษาปีที่ 3*. ปรินญานิพนธ์กศม.
(การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- นิคม กุบแก้ว. (2544). *จิตกรรมภาพทิวทัศน์ : กรณีศึกษางานจิตกรรมภาพทิวทัศน์ของ ปอล เซซานน์*.
ปรินญานิพนธ์ กศม. (ทัศนศิลป์ : ศิลปสมัยใหม่). กรุงเทพฯ : บัณฑิตวิทยาลัยมหาวิทยาลัย
ศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- นิศารัตน์ คงสวัสดิ์. (2544). *ผลการวิเคราะห์ค่าความเชื่อมั่นของผู้ตรวจแบบทดสอบความเรียงที่มีจำนวนผู้
ตรวจและวิธีการตรวจต่างกัน*. ปรินญานิพนธ์กศม.(การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย
มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- พรณี เจียมสุบุตร. (2543). *การเปรียบเทียบความเชื่อมั่นของแบบทดสอบวัดความสามารถในการ แก้โจทย์
ปัญหาทางคณิตศาสตร์ที่มีจำนวนผู้ตรวจและวิธีการตรวจต่างกัน*. ปรินญานิพนธ์ กศม. (การวัด
ผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.

- พิตร ทองชั้น. (2511). *สมรรถภาพทางสมองบางประการที่สัมพันธ์กับความสามารถทางศิลปะของนักเรียนระดับประถมศึกษาตอนปลาย*. ปรินญานิพนธ์ กศ.ม. กรุงเทพฯ : บัณฑิตวิทยาลัย วิทยาลัยวิชาการศึกษา ประสานมิตร. อัดสำเนา.
- พินดา ฟุ่งเกียรติโสข. (2540). *การประมาณค่าความเชื่อมั่นของมาตรวัดประเมินค่าที่มีรูปแบบการตอบแตกต่างกัน*. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- เพ็ญศรี พุทธิไพบูลย์. (2543). *การศึกษาค่าความเชื่อมั่นของมาตรวัดประเมินค่าแบบการเรียนรู้*. ปรินญานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร. ถ่ายเอกสาร.
- ไพรัตน์ วงษ์นาม. (2533). *สัมประสิทธิ์การอ้างอิงสรุปของแบบสอบเรียงความ*. วิทยานิพนธ์ ค.ม. (เอกวิจยและสถิติ) กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- ไพศาล หวังพานิช. (2528). *การวัดผลการศึกษา*. กรุงเทพฯ : ไทยวัฒนาพานิช.
- ภัทรา นิคมนนท์. (2532). *การประเมินผลและการสร้างแบบทดสอบ*. กรุงเทพฯ : ภาควิชาการทดสอบและวิจัยการศึกษา คณะวิชาครุศาสตร์ วิทยาลัยครูจันทระเกษม.
- . (2538). *การประเมินผลการเรียน*. กรุงเทพฯ : อักษรพิพัฒน์ จำกัด.
- บรรลือ กุลโชติ. (2530). *เกณฑ์การตัดสินการประกวดศิลปะนักเรียนระดับมัธยมศึกษาตอนต้นในโรงเรียนสังกัดกรมสามัญศึกษากรุงเทพมหานคร*. วิทยานิพนธ์ คม. (ศิลปศึกษา) จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- บุญเชิด ภิญโญนนตพงษ์. (2519). *การวัดผลและประเมินผลการศึกษา*. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- . (2521). *การวัดและประเมินผลการศึกษา : ทฤษฎีประยุกต์*. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- . (2525). *เอกสารประกอบการสอนวิชาวัดผล 301 การสร้างแบบทดสอบ 1*. มหาวิทยาลัยศรีนครินทรวิโรฒ.
- . (2527). *การทดสอบแบบอิงเกณฑ์ : แนวคิดและวิธีการ*. กรุงเทพฯ : โอเดียนสโตร์.
- . (2533). "Congeneric part Reliability." *วารสารการวัดผลการศึกษา*. 12(34)
- . (2544). *การประเมินการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ : แนวคิดและวิธีการ*. กรุงเทพฯ : สำนักงานคณะกรรมการการศึกษาแห่งชาติ.
- . (2545). "คุณภาพของเครื่องมือวัด." ใน *ประมวลสาระชุดวิชา การพัฒนาเครื่องมือสำหรับการประเมินการศึกษา* หน่วยที่ 3. กรุงเทพฯ : มหาวิทยาลัยสุโขทัยธรรมาธิราช.
- เฟียน ไชยสร. "การวัดผลงานภาคปฏิบัติ," *วารสารการวัดผลการศึกษา*. 8 (23) : 37 – 61 ; กันยายน – ธันวาคม 2529.
- มัลลิกา ยุวณะเดมิย์. (2528). *การเปรียบเทียบความเที่ยงของแบบสอบถามประเภทมาตราประเมินค่าที่มีรูปแบบคำตอบแตกต่างกัน*. วิทยานิพนธ์ ค.ม. กรุงเทพฯ : จุฬาลงกรณ์มหาวิทยาลัย, อัดสำเนา.
- วิรุณ ตั้งเจริญ. (2526). *ศิลปศึกษา*. กรุงเทพฯ : วิมลอาร์ต.
- วิราพร พงศ์อาจารย์. (2542). *การประเมินผลการเรียน*. พิษณุโลก : สถาบันราชภัฏพิบูลสงคราม.
- วิเชียร เกตุสิงห์. (2517). *การวัดผลการศึกษาและสถิติเบื้องต้น*. กรุงเทพฯ : สำนักงาน คณะกรรมการการศึกษาแห่งชาติ.
- วินัย พัฒนรัฐ. (2540). *แนวการสอนแบบเรียนมาตรฐาน ฉบับพิเศษ เน้นกระบวนการ สลน 6 ชั้นประถมศึกษาปีที่ 6*. กรุงเทพฯ : โรงพิมพ์ประสานมิตร จำกัด.
- วิญญา วิศาลาภรณ์. (2530). *การสร้างแบบทดสอบ*. มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.

- เลิศ อานันทะนะ. (2523). ศิลปะในโรงเรียน. กรุงเทพฯ : กราฟิเคอาร์ท.
- ส. วาสนา ประवालพฤกษ์. (2533). "การพิจารณาความเชื่อมั่นของแบบทดสอบการปฏิบัติ." วารสารการวัดผลการศึกษา. 12(35).
- สมนึก ภัททิยธนี. (2542). การวัดผลการศึกษา. มหาวิทยาลัยมหาสารคาม.
- สมใจ สิทธิชัย. (2535). การพัฒนาเกณฑ์ประเมินภาคปฏิบัติวิชาศิลปศึกษา หน่วยการเรียนรู้สาระระดับชั้นประถมศึกษาปีที่ 5 ตามหลักสูตรประถมศึกษา พุทธศักราช 2521 (ฉบับปรับปรุง). วิทยานิพนธ์ คม. (ศิลปศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.
- สมศักดิ์ สินธุระเวชญ์. การประเมินผลการเรียนระดับมัธยมศึกษา. กรุงเทพฯ : สำนักทดสอบทางการศึกษา กรมวิชาการ กระทรวงศึกษาธิการ, ม.ป.ป.
- สาลิกา เมฆานาวิน. (2537). การประมาณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ตามวิธีของฮวน วิธีของเบรนนอนและเคน และวิธีของราซ เมื่อกำหนดคะแนนจุดตัดต่างกัน. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุพานี กอหมั่นศิลป์. (2543). การศึกษาความเชื่อมั่นของแบบประเมินทักษะกีฬาบาสเกตบอลที่ใช้ประเมินจำนวนต่างกันประเมินนักเรียนจำนวนต่างกัน. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- สุมาลี จันทร์ชลอ. (2542). การวัดและการประเมินผล. กรุงเทพฯ : ศูนย์สื่อเสริมกรุงเทพฯ.
- สุนันท์ ศลโกสม. การประเมินผลทางการศึกษา. กรุงเทพฯ : มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร.
- สุภรณ์ ลิ้มบริบูรณ์. "การวัดผลการเรียนรู้ภาคปฏิบัติ," สารพัฒนาหลักสูตร 12 (111) กรกฎาคม – กันยายน 2535.
- เสาวนีย์ เกรียร์. (2540). คู่มือการอบรมการใช้แฟ้มสะสมผลงานนักเรียน. กรุงเทพฯ : เนติกุลการพิมพ์.
- อนันต์ ศรีโสภ. (2525). การวัดผลการศึกษา. กรุงเทพฯ : ไทยวัฒนาพานิช.
- (2524). การวัดและประเมินผลการศึกษา. กรุงเทพฯ : ไทยวัฒนาพานิช.
- อัจฉราอรุณรักษ์สมบัติ. (2542). การสร้างแบบทดสอบวัดภาคปฏิบัติโดยวิธีเขียนตอบและวิธีปฏิบัติจริง วิชางานประดิษฐ์ กลุ่มการทำงานและพื้นฐานอาชีพ ของนักเรียนชั้นประถมศึกษาปีที่ 6. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- อุษณีย์ บัวศิริพันธ์. (2543). การเปรียบเทียบค่าสัมประสิทธิ์การสุ่อ้างอิงของแบบทดสอบวิชาคณิตศาสตร์ที่มีวิธีการตรวจ จำนวนผู้ตรวจ และประสบการณ์ของผู้ตรวจ แตกต่างกัน. วิทยานิพนธ์ กศ.ม. (การวัดผลการศึกษา). กรุงเทพฯ : บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ. ถ่ายเอกสาร.
- อุบล ตูจินดา. (2533). หลักและวิธีสอนศิลปะ. กรุงเทพฯ : โอเดียนสโตร์.
- Anastasi, Anne. (1968). *Psychological Testing*. New York : Mc.Graw Hill, Inc,
- Beardsley, M.C. (1958). *Aesthetics: Problems in the Philosophy of Criticism*. New York : Harcourt, Brace & World.
- Bury-Stock, Judith A., Laurie, Cecelia, Chissom, Brad S. (1996). "Rater Agreement Indexes For Performance Assessment" *Education and Psychological measurement*. 2 : 251 – 261.
- Cohen, J.A. (1960). "Coefficient of Agreement for Nominal Scales", *Educational and Psychological Measurement*.
- Courtney, E W. (1996). "Alternative Assessment Techniques at the Elementary and Secondary Levels," เอกสารประกอบการอบรมเชิงปฏิบัติการ. คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์.

- Cleaver, Dalv G. (1966). *Problems of Value Judgement Art : An Introduction*. New York :
Harcourt, Brace & World.
- Einer, Elliot W. (1970). *Education Artistic Vision*. New York : Macmillan Publishing.
- Good, Carter Vicctor. (1973). *Dictionary of Education*. New York : McGraw - Hill.
- Gronlund, Norman E. (1976). *Measurement and Evalution in Teaching*. 3rd ed. New York :
macmillan.
- Gronlund, Norman E. (1981). *Measurement and Evalution in Teaching*. 4th ed. New York : macmillan.
- Guilford J.P. (1954). *Psychomatric Methods*. New York : Mc.Graw – Hill.
- Hopkins, C.D and Antes, R.L. (1990). *Classroom Measurement and Evaluation*. 3th Ed. Itasca, IL.
- Lindvall C. M. and Anthony J. Nitko. (1967). *Measuring Pupil Achieverment and Aptitude*.
New York : Harcourt Brace Jovanvich, Inc.
- Lord, Frederic M. and Melvin, R. Novick. (1968). *Statistical Theories of Mental Test Score*. California :
Harcourt Brace Jovanovich, Inc.
- Marshrens, John Clark and Hales, Loyde Wesley. (1971). *Classroom Test Construction*.
Massachusetts : Addison – Wesley Publishing Comparry.
- Michael, John. (1983). *Art and Adolescence : Teaching Art at the Secondary Level*. New York :
Teachers College Press, Columbia University.
- Nunnally Jum C. (1964). *Test and Measurements : Assessmeent and Predictio*. New York, Mc.Graw-
hill, Book Company.
- Ober, Richard L., Bentley, Ernest L., and Miller, Edith. (1971). *Systematic Observation of Teachng*, Prentice –
Hall, Inc., Englewood. Clifts, New Jersey.
- Saslow, C.A. (1982). *Basic Research- Methods*. Addison – Wesley Publishing Co.
- Stenmark, J.K. (1991). *Mathematics Assessment : myths, models, good question, and Practical
Suggestion*. University of California : The Nation Council of Teachers of Mathematics, INC.
- Simpson, E.J. (1966). *The Classification of Educational Objective : Psychomotor Domain*. Illinois :
University of Illinois, Urbana – Champaign.
- Tasse, Marc J. and others. (1994). "Test-retest and Inter-rater Reliability of the Quebec Adaptive
Behavior Scale." *NISC DISC REPORT*. _ERIC 1985-1999.
- Thorndike, Robert L. and Hagen Elizabeth. (1960). *Measurement and Evaluation in Psycholgy and
Education*. New york : Willey.
- Tuckman, Bruce W. (1975). *Measuring Education Outcomes Fundamentals of Testing*. New York :
Harcourt Bruce Jovanovich, Inc.
- Turney Billy L. and Robb George P. (1971). *Research in Educational and Introduction*. Illinois : The
Dryden Press Inc.
- Uebersax, John. (2000). "Raw Agreement Indices," <http://ourworld.compuserve.com>.
- Vockeii, E.P. and William, J.A. (1995). *Education Research*. New Jersey : Prentice – Hall Inc.
- William, M.K. (2001). "Raw Agreement Indices," <http://ourworld.compuserve.com>.

ภาคผนวก

ภาคผนวก ก

1. รายนามผู้เชี่ยวชาญในการตรวจสอบคุณภาพเครื่องมือ
2. ดัชนีความสอดคล้องของผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือ (*IOC*)
3. ดัชนีค่าอำนาจจำแนกของการประเมินผลงานศิลปะ
4. คู่มือการประเมินผลงานศิลปะแบบเกณฑ์แยก (*Analytic Scoring Rubric*)
5. คู่มือการประเมินผลงานศิลปะแบบเกณฑ์รวม (*Holistic Scoring Rubric*)

รายนามผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือ

ผู้วิจัยได้เลือกผู้เชี่ยวชาญที่มีวุฒิการศึกษาตั้งแต่ระดับปริญญาโท สาขาเอกการวัดผลการศึกษา และครูผู้สอน ที่มีวุฒิปริญญาตรี วิชาเอกศิลปศึกษา และมีประสบการณ์การสอนวิชาศิลปศึกษามาเป็นเวลา 5 ปี ขึ้นไป ซึ่งมีรายนามผู้เชี่ยวชาญในการตรวจสอบคุณภาพเครื่องมือ ดังนี้

1. ผู้เชี่ยวชาญทางการวัดผล

- | | |
|----------------------------|---|
| 1. นายชาญวิทย์ จรัสสุทธิธร | ศึกษานิเทศก์ เขตการศึกษา 11 |
| 2. นางนิศารัตน์ คงสวัสดิ์ | หัวหน้างานวัดผล โรงเรียนอัสสัมชัญแผนกประถม
กรุงเทพมหานคร |

2. ผู้เชี่ยวชาญทางการสอนวิชาศิลปศึกษา

- | | |
|-------------------------|---|
| 3. นายไพฑูรย์ ภูมิซ้อ | ผู้ช่วยอาจารย์ใหญ่โรงเรียนสุเหร่าคลองหนึ่ง |
| 4. นางสาวปนัดดา ชื่นเชย | อาจารย์สอนวิชาศิลปศึกษาโรงเรียนบ้านหนองระแหง |
| 5. นายธงชัย ดวงมุลตรี | อาจารย์สอนวิชาศิลปศึกษาโรงเรียนสุเหร่าคลองหนึ่ง |

ตาราง 12 ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือ (*IOC*) ของการให้คะแนน
การประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric)

ข้อ	ผู้เชี่ยวชาญ					ค่า <i>IOC</i>	ผลการคัดเลือก
	1	2	3	4	5		
1	1	1	1	1	1	1.0	ใช้ได้
2	1	1	1	1	1	1.0	ใช้ได้
3	1	1	1	1	1	1.0	ใช้ได้
4	1	1	1	1	1	1.0	ใช้ได้
5	1	1	1	1	1	1.0	ใช้ได้
6	1	1	1	1	1	1.0	ใช้ได้
7	1	1	1	1	1	1.0	ใช้ได้

ตาราง 13 ค่าดัชนีความสอดคล้องของผู้เชี่ยวชาญตรวจสอบคุณภาพเครื่องมือ (*IOC*) ของการให้คะแนน
การประเมินแบบเกณฑ์รวม (Holistic Scoring Rubric)

ข้อ	ผู้เชี่ยวชาญ					ค่า <i>IOC</i>	ผลการคัดเลือก
	1	2	3	4	5		
1	1	1	1	1	1	1.0	ใช้ได้

ตาราง 14 ค่าอำนาจจำแนก (r_{xy}) ของแบบประเมินผลงานศิลปะ

ภาพที่	ค่าอำนาจจำแนก (r_{xy}) ของ การให้คะแนนแบบเกณฑ์แยก	ค่าอำนาจจำแนก (r_{xy}) ของ การให้คะแนนแบบเกณฑ์รวม	ผลการทดลอง
1	0.999	0.999	ใช้ได้
2	1.000	0.999	ใช้ได้
3	0.999	0.999	ใช้ได้
4	0.999	0.999	ใช้ได้
5	0.999	0.999	ใช้ได้
6	0.999	0.999	ใช้ได้
7	0.999	0.999	ใช้ได้
8	0.999	0.999	ใช้ได้
9	0.999	1.000	ใช้ได้
10	0.999	0.999	ใช้ได้

คู่มือการประเมินผลงานศิลปะ
การประเมินผลงานศิลปะแบบเกณฑ์แยก (Analytic Scoring Rubric)

คำชี้แจง

คู่มือประเมินนี้ประกอบด้วย

- แบบประเมินผลงานศิลปะ จำนวน 7 ข้อ
- กฎเกณฑ์มาตรฐานประเมินแบบเกณฑ์แยก

ขั้นตอนการประเมิน ดำเนินการดังนี้

1. ให้คุณครูดูภาพศิลปะที่กำหนดให้ทีละ 1 ภาพ
2. ใช้คู่มือการประเมินนี้เป็นแนวทางในการประเมินผลงานศิลปะของนักเรียน โดยพิจารณาว่าผลงานศิลปะของนักเรียน มีประเด็นตามกฎเกณฑ์มาตรฐานประเมินหรือไม่ มากน้อยเพียงใด
3. ประเมินโดยแบบเกณฑ์แยก (Analytic Scoring Rubrics) ซึ่งมีขั้นตอนการประเมิน ดังนี้

3.1 พิจารณาผลงานศิลปะตามกฎเกณฑ์การให้คะแนนที่กำหนดให้ทีละภาพ

จนครบ 30 ภาพ

กฎเกณฑ์การให้คะแนนการประเมินภาพระบายสีแบบเกณฑ์แยก (Analytic Scoring Rubric)

ประเด็นประเมิน	คำอธิบายระดับคุณภาพ			
	4	3	2	1
1. การเน้นจุดเด่นของภาพ	- ภาพเด่นสะดุดตาและมีการใช้สีสะดุดตาชัดเจน	- ภาพเด่นสะดุดตาแต่การใช้สีกระจายจนไม่มีจุดเด่น	- มีการใช้สีมากหรือน้อยไปไม่เหมาะสมกับภาพ	- ภาพไม่ชัดเจนและการให้สีน้อยมาก
2. ความสมดุลของภาพ	- มีการจัดวางภาพเหมาะสมกับหน้ากระดาษและภาพมีระยะใกล้-ไกล มีการผสมผสานกลมกลืนของรูปและพื้นหลัง	- มีการจัดวางภาพเหมาะสมกับหน้ากระดาษ และมีระยะใกล้ - ไกลของภาพ	- มีการจัดวางภาพเหมาะสมกับหน้ากระดาษแต่ขาดระยะใกล้ - ไกลของภาพ	- ภาพไม่มีลักษณะของการวางภาพที่มีความสมดุลเลย
3. การระบายสี	- มีการระบายสีภาพอยู่ในกรอบสะอาดเรียบร้อย สีเรียบ กลมกลืน	- มีการระบายสีภาพออกนอกเส้นเป็นส่วนน้อย สีส่วนใหญ่เรียบ	- มีการระบายสีภาพออกนอกกรอบเลอะเทอะ สีไม่เรียบ	- มีการระบายสีภาพอย่างไม่มีจุดมุ่งหมาย
4. การใช้สี	- ภาพให้แสงและเงาถูกต้อง และแยกความสัมพันธ์ภาพและพื้นหลังชัดเจน ให้สีเหมือนธรรมชาติ	- ภาพให้แสงและเงาแต่ไม่ถูกต้อง และแยกความสัมพันธ์ภาพและพื้นหลังชัด มีการให้สีใกล้เคียงธรรมชาติ	- ภาพไม่แสดงแสงและเงา แต่ในภาพมีสีพื้นหลังกลมกลืนมีการให้สีผิดธรรมชาติ	- ภาพไม่แสดงแสงและเงา มีสีพื้นหลังกลมกลืนไม่เด่นชัด มีการให้สีผิดธรรมชาติ
5. ความคิดสร้างสรรค์	- ภาพมีแนวคิดแปลกใหม่ มีการสื่อความหมายของการใช้ความคิดของตัวเองและไม่ซ้ำภาพผู้อื่น	- ภาพมีแนวคิดแปลกใหม่ มีการสื่อความหมายของการใช้ความคิดของตัวเอง แต่ซ้ำภาพผู้อื่น	- ภาพมีแนวคิดแปลกใหม่ ไม่มีการสื่อความหมายของการใช้ความคิดของตัวเอง	- ภาพไม่มีแนวคิดแปลกใหม่และซ้ำกับภาพผู้อื่น
6. ความประณีต	- ภาพมีความละเอียดมีความชัดเจนของภาพและสามารถมองเห็นความงดงามของภาพ	- ภาพมีความละเอียดมีความชัดเจนของภาพแต่ขาดความงดงาม	- ภาพมีความละเอียดแต่ไม่มีความชัดเจนของภาพ	- ไม่มีความละเอียดของภาพและไม่มี ความชัดเจนของภาพ
7. ความสะอาด	- ภาพมีการระบายสีอยู่ในกรอบไม่เลอะเทอะและมีความเป็นระเบียบเรียบร้อย	- ภาพมีการระบายสีอยู่ในกรอบไม่เลอะเทอะแต่ขาดความเป็นระเบียบเรียบร้อย	- ภาพมีการระบายสีอยู่ในกรอบ เลอะเทอะและไม่มีความเป็นระเบียบ	- ภาพมีการระบายสีออกนอกกรอบเลอะเทอะ

คู่มือการประเมินผลงานศิลปะ
แบบประเมินผลงานศิลปะแบบเกณฑ์รวม (Holistic Scoring Rubric)

คำชี้แจง

คู่มือประเมินนี้ประกอบด้วย

- แบบประเมินผลงานศิลปะ
- กฎเกณฑ์มาตรฐานประเมินแบบเกณฑ์รวม

ขั้นตอนการประเมิน ดำเนินการดังนี้

1. ให้คุณครูดูภาพศิลปะที่กำหนดให้ทีละ 1 ภาพ
2. ใช้คู่มือการประเมินนี้เป็นแนวทางในการประเมินผลงานศิลปะของนักเรียน โดยพิจารณาว่าผลงานศิลปะของนักเรียน มีประเด็นตามกฎเกณฑ์มาตรฐานประเมินหรือไม่ มากน้อยเพียงใด

3. (ดำเนินการหลังการประเมินแบบเกณฑ์แยกเสร็จแล้ว 1 สัปดาห์)
 ประเมินโดยแบบเกณฑ์รวม (Holistic Scoring Rubrics) ซึ่งมีขั้นตอนการประเมิน ดังนี้

3.1 พิจารณาผลงานศิลปะตามกฎเกณฑ์การให้คะแนนที่กำหนดให้ทีละภาพ

จนครบ 30 ภาพ

กฎเกณฑ์การให้คะแนนการประเมินภาพระบายสีแบบเกณฑ์รวม (Holistic Scoring Rubrics)

ระดับคุณภาพ	คำอธิบาย
4	- ภาพเด่นสะดุดตา มีการจัดวางได้เหมาะสมกับกระดาษ แสดงระยะใกล้-ไกล การระบายสีกลมกลืน ให้แสงและเงาถูกต้องเหมือนธรรมชาติ และมีแนวคิดแปลกใหม่ของเป็นตัวเอง สามารถมองเห็นได้ชัดเจนงดงามเป็นระเบียบ
3	- ภาพเด่นสะดุดตา จัดวางได้เหมาะสมกับกระดาษ ระบายสีออกนอกเส้นบางแต่อยู่ในกรอบ และ ให้แสงและเงาไม่ถูกต้อง ให้สีใกล้เคียงธรรมชาติ มีความแปลกใหม่ สื่อความหมายของตัวเอง มีความชัดเจนของภาพ
2	- ภาพมีการใช้สีมากไม่เหมาะสม จัดวางภาพเหมาะสมกับกระดาษ ไม่มีระยะใกล้-ไกล ระบายสีเลอะเทอะ ไม่แสดงแสงและเงา ให้สีผิดธรรมชาติ มีแนวคิดแปลกใหม่ แต่ไม่มีความชัดเจน ไม่มีความเป็นระเบียบของภาพ
1	- ภาพไม่ชัดเจน การวางภาพไม่มีความสมดุล มีการระบายสีกระจาย เลอะเทอะทั่วไป ไม่แสดงแสงและเงา ให้สีผิดธรรมชาติ ไม่มีแนวคิดของตัวเอง ไม่มีความละเอียดของภาพ ไม่มีความสะอาด

แบบประเมินผลงานศิลปะ ของแบบประเมินแบบเกณฑ์แยก (Analytic Scoring Rubric)

คำชี้แจง โปรดใส่เครื่องหมาย ✓ ลงในช่องคะแนนที่ตรงกับการประเมินผลงานศิลปะของท่าน โดยเทียบกับกฎเกณฑ์การให้คะแนนที่กำหนด

ภาพที่...	ประเด็นประเมิน	ระดับคุณภาพคะแนน			
		4	3	2	1
1. การเน้นจุดเด่นของภาพ					
2. ความสมดุลของภาพ					
3. การระบายสี					
4. การใช้สี					
5. ความคิดสร้างสรรค์					
6. ความประณีต					
7. ความสะอาด					
คะแนนรวม					

ภาพที่....	ประเด็นประเมิน	ระดับคุณภาพคะแนน			
		4	3	2	1
1. การเน้นจุดเด่นของภาพ					
2. ความสมดุลของภาพ					
3. การระบายสี					
4. การใช้สี					
5. ความคิดสร้างสรรค์					
6. ความประณีต					
7. ความสะอาด					
คะแนนรวม					

ภาพที่...	ประเด็นประเมิน	ระดับคุณภาพคะแนน			
		4	3	2	1
1. การเน้นจุดเด่นของภาพ					
2. ความสมดุลของภาพ					
3. การระบายสี					
4. การใช้สี					
5. ความคิดสร้างสรรค์					
6. ความประณีต					
7. ความสะอาด					
คะแนนรวม					

ภาพที่....	ประเด็นประเมิน	ระดับคุณภาพคะแนน			
		4	3	2	1
1. การเน้นจุดเด่นของภาพ					
2. ความสมดุลของภาพ					
3. การระบายสี					
4. การใช้สี					
5. ความคิดสร้างสรรค์					
6. ความประณีต					
7. ความสะอาด					
คะแนนรวม					

เกณฑ์การให้คะแนน

คะแนน 1 – 7 คะแนน หมายถึง 1 คะแนน

คะแนน 8 – 14 คะแนน หมายถึง 2 คะแนน

คะแนน 15 – 21 คะแนน หมายถึง 3 คะแนน

คะแนน 22 – 28 คะแนน หมายถึง 4 คะแนน

ชื่อผู้ประเมิน โรงเรียน

สำนักงานเขตคลองสามวา กรุงเทพมหานคร

แบบประเมินผลงานศิลปะแบบเกณฑ์รวม (Holistic Score Rubric)

คำชี้แจง โปรดใส่เครื่องหมาย ✓ ลงในช่องคะแนนที่ตรงกับการประเมินผลงานของท่าน โดยเทียบกับกฎเกณฑ์การให้คะแนนที่กำหนด

ประเด็นประเมิน	ภาพที่	ระดับคุณภาพคะแนน			
		4	3	2	1
• การเน้นจุดเด่นของภาพ • ความสมดุลของภาพ • การระบายสี • การใช้สี • ความคิดสร้างสรรค์ • ความประณีต • ความสะอาด	1				
	2				
	3				
	4				
	5				
	6				
	7				
	8				
	9				
	10				
	11				
	12				
	13				
	14				
	15				
	16				
	17				
	18				
	19				
	20				
	21				
	22				
	23				
	24				
	25				
	26				
	27				
	28				
	29				
	30				

ภาคผนวก ข

1. รายนามผู้ประเมินผลงานศิลปะ
2. เปรียบเทียบค่าสถิติพื้นฐาน การประเมินด้วยการให้คะแนนการประเมินแบบเกณฑ์แยก และแบบเกณฑ์รวม
3. คะแนนการประเมินผลงานศิลปะ เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก และค่าสถิติพื้นฐาน
4. ผลการประเมินแบบประเมินผลงานศิลปะ เมื่อใช้การประเมินแบบเกณฑ์รวม และค่าสถิติพื้นฐาน
5. ค่าความเชื่อมั่นสูตร P_0 ของแฮมเบิลตันและโนวิก
เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก
6. ค่าความเชื่อมั่นสูตร P_0 ของแฮมเบิลตันและโนวิก
เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวม
7. ค่าความเชื่อมั่นสูตร K ของสก็อต เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก
8. ค่าความเชื่อมั่นสูตร K ของสก็อต เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวม
9. ค่าความเชื่อมั่นสูตร RAI ของเบอร์รี่ - สต็อก และคณะ
เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก
10. ค่าความเชื่อมั่นสูตร RAI ของเบอร์รี่ - สต็อก และคณะ
เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวม

รายนามผู้ตรวจแบบประเมินผลงานศิลปะ

ชื่อ นามสกุล นายสุเทพ อินทกุล(R1)
 ตำแหน่ง อาจารย์ 1 ระดับ 5 อาจารย์ผู้สอนวิชาศิลปศึกษา
 สถานที่ทำงาน โรงเรียนประชาราษฎร์อุปถัมภ์วิทยา
 วุฒิต่างการศึกษา คบ. ศิลปศึกษา
 ประสบการณ์สอน 10 ปี

ชื่อ นามสกุล นายวิทยา ตันทัยย์(R2)
 ตำแหน่ง อาจารย์ 2 ระดับ 7 อาจารย์ผู้สอนวิชาศิลปศึกษา
 สถานที่ทำงาน โรงเรียนสุเหร่าคลองหนึ่ง
 วุฒิต่างการศึกษา คบ. ศิลปศึกษา
 ประสบการณ์สอน 25 ปี

ชื่อ นามสกุล นายมานพ มินยา(R3)
 ตำแหน่ง อาจารย์ 2 ระดับ 6 อาจารย์ผู้สอนวิชาศิลปศึกษา
 สถานที่ทำงาน โรงเรียนสุเหร่าสามวา
 วุฒิต่างการศึกษา คบ. ศิลปศึกษา
 ประสบการณ์สอน 20 ปี

ชื่อ นามสกุล นายภัทรพงศ์ สุดตะ(R4)
 ตำแหน่ง อาจารย์ 2 ระดับ 7 อาจารย์ผู้สอนวิชาศิลปศึกษา
 สถานที่ทำงาน โรงเรียนสุเหร่าคลองสี่
 วุฒิต่างการศึกษา คบ. ศิลปศึกษา
 ประสบการณ์สอน 25 ปี

ชื่อ นามสกุล นายผดุงศักดิ์ ศัพทเสน(R5)
 ตำแหน่ง อาจารย์ 2 ระดับ 7 อาจารย์ผู้สอนวิชาศิลปศึกษา
 สถานที่ทำงาน โรงเรียนวัดสุทธิสะอาด
 วุฒิต่างการศึกษา คบ. ศิลปศึกษา
 ประสบการณ์สอน 25 ปี

ตาราง 15 เปรียบเทียบค่าสถิติพื้นฐาน การประเมินด้วยการให้คะแนนการประเมินแบบเกณฑ์แยก
และแบบเกณฑ์รวม

ผู้ประเมิน	การให้คะแนนการประเมินแบบเกณฑ์แยก			การให้คะแนนการประเมินแบบเกณฑ์รวม		
	$\bar{X}$	<i>S.D.</i>	<i>C.V.</i>	$\bar{X}$	<i>S.D.</i>	<i>C.V.</i>
คนที่ 1	3.00	0.186	6.187	2.700	0.904	33.474
คนที่ 2	3.467	0.507	14.629	2.333	0.607	25.997
คนที่ 3	3.100	0.421	13.565	2.767	0.679	24.536
คนที่ 4	3.233	0.430	13.306	3.300	0.466	14.124
คนที่ 5	3.000	0.587	19.573	2.800	0.610	21.796
เฉลี่ยรวม	3.160	0.426	13.452	2.780	0.653	23.985

ตาราง 16 คะแนนการประเมินผลงานศิลปะ เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก และค่าสถิติพื้นฐาน

ภาพที่	เต็ม	R1	R2	R3	R4	R5	ค่าเฉลี่ย
1	4	3	4	3	3	3	3.20
2	4	3	4	4	3	3	3.40
3	4	3	4	3	4	2	3.20
4	4	3	3	3	3	2	2.80
5	4	3	4	3	4	3	3.40
6	4	3	3	3	3	2	2.80
7	4	3	4	3	3	3	3.20
8	4	3	3	3	3	2	2.80
9	4	3	3	3	3	3	3.00
10	4	3	3	3	3	2	2.80
11	4	3	3	4	4	3	3.40
12	4	3	4	3	3	3	3.20
13	4	3	3	3	3	3	3.00
14	4	2	3	3	3	3	2.80
15	4	3	4	3	3	3	3.20
16	4	3	3	3	3	3	3.00
17	4	3	4	3	3	3	3.20
18	4	3	4	3	3	3	3.20
19	4	3	3	3	3	3	3.00
20	4	3	3	3	3	4	3.20
21	4	3	3	3	4	3	3.20
22	4	3	3	3	4	3	3.20
23	4	3	3	3	3	3	3.00
24	4	3	3	3	3	3	3.00
25	4	3	4	3	3	4	3.40
26	4	3	3	3	4	3	3.20
27	4	3	4	3	4	4	3.60
28	4	3	4	3	3	4	3.40
29	4	3	4	4	3	3	3.40
30	4	4	4	3	3	4	3.60

ตาราง 16 (ต่อ)

ค่าเฉลี่ย	3.00	3.46	3.10	3.23	3.00	3.15
<i>S.D.</i>	0.18	0.50	0.42	0.43	0.58	0.42
<i>C.V.</i>	6.18	14.62	13.56	13.30	19.57	13.44

ตาราง 17 คะแนนการประเมินผลงานศิลปะ เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวม และค่าสถิติพื้นฐาน

ภาพที่	เต็ม	R1	R2	R3	R4	R5	ค่าเฉลี่ย
1	4	3	2	4	3	2	2.80
2	4	3	3	3	3	2	2.80
3	4	3	3	3	4	2	3.00
4	4	2	2	2	4	2	2.40
5	4	2	3	2	3	2	2.40
6	4	2	2	3	4	2	2.60
7	4	3	3	3	3	3	3.00
8	4	2	2	2	4	2	2.40
9	4	3	2	3	3	3	3.20
10	4	2	1	2	4	2	2.20
11	4	3	2	3	3	3	2.80
12	4	3	3	3	3	3	3.00
13	4	3	2	3	3	3	2.80
14	4	2	3	2	3	2	2.40
15	4	3	3	3	3	3	3.00
16	4	3	2	3	3	3	2.80
17	4	2	3	4	4	3	3.20
18	4	2	2	3	3	3	2.60
19	4	2	1	2	3	3	2.20
20	4	3	3	2	4	4	3.20
21	4	3	2	2	4	3	2.80
22	4	3	3	2	3	3	2.80
23	4	3	2	3	3	3	2.80
24	4	2	2	2	3	3	2.40
25	4	3	2	2	4	4	3.00
26	4	3	2	3	3	3	2.80
27	4	3	2	3	3	3	2.80
28	4	3	2	4	3	4	3.00
29	4	3	3	4	3	3	3.20
30	4	4	3	3	3	3	3.20

ตาราง 17 (ต่อ)

ค่าเฉลี่ย	2.70	2.33	2.76	3.30	2.80	2.78
<i>S.D.</i>	0.90	0.60	0.67	0.46	0.61	0.65
<i>C.V.</i>	33.47	25.99	24.53	14.12	21.79	23.98

ตาราง 18 ค่าความเชื่อมั่นของสูตร P_0 ของแฮมเบิลตัน และโนวิก เมื่อใช้การให้คะแนนการประเมิน
แบบเกณฑ์แยก

ผู้ประเมิน	ประเมินครั้งที่		P_{11}	P_{22}	$P_0 = P_{11} + P_{22}$
คนที่ 1	1	2	0.970	0.000	0.970
คนที่ 2	1	2	0.930	0.000	0.930
คนที่ 3	1	2	1.000	0.000	1.000
คนที่ 4	1	2	1.000	0.000	1.000
คนที่ 5	1	2	0.530	0.100	0.630
ค่าเฉลี่ยรวม					0.906

ตาราง 19 ค่าความเชื่อมั่นของสูตร P_0 ของแฮมเบิลตัน และโนวิก เมื่อใช้การให้คะแนนการ
แบบเกณฑ์รวม

ผู้ประเมิน	ประเมินครั้งที่		P_{11}	P_{22}	$P_0 = P_{11} + P_{22}$
คนที่ 1	1	2	0.670	0.200	0.870
คนที่ 2	1	2	0.300	.0170	0.470
คนที่ 3	1	2	0.630	0.000	0.630
คนที่ 4	1	2	1.000	0.000	1.000
คนที่ 5	1	2	0.430	0.100	0.530
ค่าเฉลี่ยรวม					0.700

ตาราง 20 ค่าความเชื่อมั่นของสูตร K ของสก็อต เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก

ผู้ประเมิน	P_0	P_e	$K = p_0 - P_e / 1 - P_e$
คนที่ 1 คนที่ 2	0.853	0.003	0.853
คนที่ 1 คนที่ 3	0.921	0.002	0.922
คนที่ 1 คนที่ 4	0.872	0.002	0.872
คนที่ 1 คนที่ 5	0.891	0.003	0.891
คนที่ 2 คนที่ 3	0.851	0.002	0.851
คนที่ 2 คนที่ 4	0.834	0.002	0.834
คนที่ 2 คนที่ 5	0.820	0.003	0.819
คนที่ 3 คนที่ 4	0.892	0.002	0.892
คนที่ 3 คนที่ 5	0.846	0.002	0.846
คนที่ 4 คนที่ 5	0.819	0.003	0.819
ค่าเฉลี่ยรวม			0.864

ตาราง 21 ค่าความเชื่อมั่นของสูตร K ของสก็อต เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวม

ผู้ประเมิน	P_0	P_e	$K = p_0 - P_e / 1 - P_e$
คนที่ 1 คนที่ 2	0.759	0.002	0.758
คนที่ 1 คนที่ 3	0.840	0.003	0.839
คนที่ 1 คนที่ 4	0.759	0.002	0.759
คนที่ 1 คนที่ 5	0.848	0.003	0.848
คนที่ 2 คนที่ 3	0.724	0.003	0.724
คนที่ 2 คนที่ 4	0.759	0.002	0.759
คนที่ 2 คนที่ 5	0.709	0.003	0.709
คนที่ 3 คนที่ 4	0.725	0.003	0.724
คนที่ 3 คนที่ 5	0.813	0.003	0.812
คนที่ 4 คนที่ 5	0.763	0.002	0.762
ค่าเฉลี่ยรวม			0.769

ตาราง 22 ค่าความเชื่อมั่นของสูตร RAI ของเบอร์รี่ – สตัดอก และคณะ เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์แยก

ผู้ประเมิน	$ R_m - \bar{R} $	$\sum^M R_m - \bar{R} $	RAI
คนที่ 1	7.000	45.600	0.873
คนที่ 2	11.800		
คนที่ 3	7.200		
คนที่ 4	9.200		
คนที่ 5	10.400		

ตาราง 23 ค่าความเชื่อมั่นของสูตร RAI ของเบอร์รี่ – สตัดอก และคณะ เมื่อใช้การให้คะแนนการประเมินแบบเกณฑ์รวม

ผู้ประเมิน	$ R_m - \bar{R} $	$\sum^M R_m - \bar{R} $	RAI
คนที่ 1	8.400	66.000	0.817
คนที่ 2	16.600		
คนที่ 3	12.400		
คนที่ 4	16.800		
คนที่ 5	11.800		

ประวัติผู้วิจัย

ประวัติย่อผู้วิจัย

ชื่อ ชื่อสกุล	นายชัยยุทธ คชรินทร์
วันเดือนปีเกิด	12 มีนาคม 2516
สถานที่เกิด	อ. ท่าแซะ จ. ชุมพร
สถานที่อยู่ปัจจุบัน	5/257 ซ. รามอินทรา 14 แขวงจรเข้บัว เขตลาดพร้าว กรุงเทพมหานคร 10203
ตำแหน่งหน้าที่การงานในปัจจุบัน	อาจารย์ 1 ระดับ 4
สถานที่ทำงานปัจจุบัน	โรงเรียนสุเหร่าคลองหนึ่ง 15/6 หมู่ 11 ถนนสุเหร่าคลองหนึ่ง แขวงบางชัน เขตคลองสามวา กรุงเทพมหานคร 10510
ประวัติการศึกษา	
พ.ศ. 2532	มัธยมศึกษาตอนต้น โรงเรียนท่าแซะรัชดาภิเษก จ. ชุมพร
พ.ศ. 2535	มัธยมศึกษาตอนปลาย โรงเรียนท่าแซะรัชดาภิเษก จ. ชุมพร
พ.ศ. 2539	คบ. เอกวิทยาศาสตร์ สถาบันราชภัฏจันทรเกษม
พ.ศ. 2546	กศ.ม. สาขาวิชาการวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ