


การประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง  
ของนักเรียนโรงเรียนในสังกัดเมืองพัทยาโดยบูรณาการรูปแบบ  
การประเมินของเครื่องแพตทริกและการประเมินอภิमान


คำป๋น ศรีมหาไชย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาวิจัย วัฒน และสถิติการศึกษา  
คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา  
กุมภาพันธ์ 2560  
ลิขสิทธิ์เป็นของมหาวิทยาลัยบูรพา


คณะกรรมการควบคุมวิทยานิพนธ์และคณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณา  
วิทยานิพนธ์ของ คำปิ่น ศรีมหาไชย ฉบับนี้แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตาม  
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัฒนผลและสถิติการศึกษา ของมหาวิทยาลัยบูรพาได้


คณะกรรมการควบคุมวิทยานิพนธ์

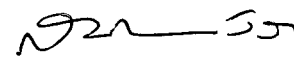
  
.....อาจารย์ที่ปรึกษาหลัก  
(ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร)

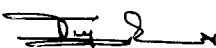
  
.....อาจารย์ที่ปรึกษาร่วม  
(ดร.สมศักดิ์ ลิลา)

คณะกรรมการสอบวิทยานิพนธ์


  
.....ประธาน  
(ผู้ช่วยศาสตราจารย์ ดร.จตุมา รัตนพลแสนย์)

  
.....กรรมการ  
(ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร)

  
..... กรรมการ  
(ดร.สมศักดิ์ ลิลา)

  
.....กรรมการ  
(ดร.เสกสรรค์ ทองคำบรรจง)

คณะศึกษาศาสตร์อนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตาม  
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัฒนผลและสถิติการศึกษา ของมหาวิทยาลัยบูรพา

  
..... คณบดีคณะศึกษาศาสตร์  
(รองศาสตราจารย์ ดร.วิจิต สุรัตน์เรืองชัย)

วันที่ 6 เดือน มีนาคม พ.ศ. 2560

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้ด้วยความกรุณาจาก ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร อาจารย์ที่ปรึกษาหลัก ดร.สมศักดิ์ ติลา อาจารย์ที่ปรึกษาร่วม ที่กรุณาให้คำปรึกษาแนะนำแนวทางที่ถูกต้อง ตลอดจนแก้ไขข้อบกพร่องต่าง ๆ ด้วยความละเอียดถี่ถ้วน และเอาใจใส่ด้วยดีเสมอมา ผู้วิจัยรู้สึกซาบซึ้งเป็นอย่างยิ่ง จึงขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.จตุมา รัตนพลแสนย์ อาจารย์ประจำ คณะครุศาสตร์ มหาวิทยาลัยราชภัฏบ้านสมเด็จเจ้าพระยา ผู้ทรงคุณวุฒิภายนอกที่กรุณาให้ความรู้ให้คำปรึกษา ตรวจสอบแก้ไขและวิจารณ์ผลงานทำให้งานมีความสมบูรณ์ยิ่งขึ้น และผู้ทรงคุณวุฒิทุกท่านที่ให้ความอนุเคราะห์ในการตรวจสอบรวมทั้งให้คำแนะนำแก้ไขเครื่องมือที่ใช้ในการวิจัยให้มีคุณภาพ นอกจากนี้ยังได้รับความอนุเคราะห์จากหัวหน้าโครงการ เจ้าหน้าที่ ตลอดจนผู้เข้ารับการอบรมในโครงการที่ให้ความร่วมมือเป็นอย่างดีในการเก็บรวบรวมข้อมูลที่ใช้ในการวิจัยทำให้วิทยานิพนธ์ฉบับนี้สำเร็จได้ด้วยดี

ขอกราบขอบพระคุณ คุณพ่อทองปลิว คุณแม่วร ศรีมหาไชย พี่และน้อง ๆ ทุกคนที่ให้กำลังใจ และสนับสนุนผู้วิจัยเสมอมา

คุณค่าและประโยชน์ของวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นกตัญญูคุณเวทิตาแด่ บุษปาริ บุรพาจารย์ และผู้มีพระคุณทุกท่านทั้งในอดีตและปัจจุบัน ที่ทำให้ข้าพเจ้าเป็นผู้มีการศึกษา และประสบความสำเร็จมาจนตราบนานเท่านานนี้

คำปิ่น ศรีมหาไชย

55910200: สาขาวิชา: วิชา วัตถุประสงค์ และสถิติการศึกษา; วท.ม. (วิชา วัตถุประสงค์ และสถิติการศึกษา)

คำสำคัญ: ประเมินโครงการ/ ข้อเสนอมาตรฐาน/ บูรณาการรูปแบบ

คำป็น ศรีมหาไชย: การประเมินโครงการสร้างข้อเสนอมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิमान (THE EVALUATION OF HIGHER LEVEL THINKING EVALUATION OF PATTAYA STUDENTS BY INTEGRATING KIRKPATRICK EVALUATION MODEL OF AND META EVALUATION APPROACH) คณะกรรมการควบคุมวิทยานิพนธ์: พงศ์เทพ จิระโร, ค.ค., สมศักดิ์ ธิลา, กศ.ค., จุติมา รัตนพลแสนย์, ปร.ค. 130 หน้า. ปี พ.ศ. 2560.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อประเมินโครงการสร้างข้อเสนอมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิमान โดยมีวัตถุประสงค์เฉพาะ เพื่อประเมินปฏิกิริยาตอบสนอง (Reaction evaluation) การเรียนรู้ (Learning evaluation) พฤติกรรม (Behavior evaluation) ผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) โครงการตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิमान กลุ่มตัวอย่างบุคลากรที่เข้าอบรมในโครงการสร้างข้อเสนอมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของเมืองพัทยา จำนวน 114 คน เครื่องมือที่ใช้แบบประเมินโครงการ สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน สถิติการทดสอบค่าที สถิติการวิเคราะห์ความแปรปรวนแบบทางเดียว การวิเคราะห์สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน และการวิเคราะห์เนื้อหา

ผลการวิจัยปรากฏว่า ด้านปฏิกิริยาตอบสนองผู้เข้าอบรมมีความพึงพอใจในสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในการฝึกอบรมเหมาะสมเพียงพอพร้อมทั้งมีความคิดเห็นว่าควรปรับระยะเวลาในการอบรมให้กระชับ ด้านการเรียนรู้ผู้เข้าอบรมมีความพึงพอใจในความรู้ ความเข้าใจ การสร้างข้อสอบมากยิ่งขึ้นเพราะในเนื้อหามีการนำสาระแกนกลางมาใช้ออกข้อสอบในแต่ละตัวชี้วัดด้วย ด้านพฤติกรรมผู้เข้ารับการอบรมสามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการจัดการเรียนการสอนรวมถึงการปฏิบัติงานและมีความเป็นไปได้ว่าข้อสอบที่สร้างมีคุณภาพและสามารถวัดผู้เรียนได้ตรงตามมาตรฐานได้ และสุดท้ายด้านผลลัพธ์ที่เกิดต่อองค์กรผู้เข้าอบรมมีความพึงพอใจในการปฏิรูประบบการทำงานเพื่อนำไปใช้ในการปฏิบัติงานจริงและประโยชน์ที่ได้รับมีความรู้ เทคนิค และขั้นตอนการออกข้อสอบซึ่งสามารถนำไปใช้ได้จริง

55910200: MAJOR: EDUCATIONAL RESEARCH, MEASUREMENT AND STATISTICS;  
M.SC. (EDUCATIONAL RESEARCH, MEASUREMENT AND STATISTICS)

KEYWORD: EVALUATION PROJECT/ STANDARD TEST/ INTEGRATED  
EVALUATION MODEL

KHAMPAN SRIMAHACHAI: THE EVALUATION OF HIGHER LEVEL  
THINKING EVALUATION OF PATTAYA STUDENTS BY INTEGRATING  
KIRKPATRICK EVALUATION MODEL OF AND META EVALUATION APPROACH.  
ADVISORY COMMITTEE: PONGTHEP JIRARO, Ed.D., SOMSAK LILA, Ed.D., JITIMA  
RATTANAPOLSAEN, Ph.D. 130 P. 2017.

The objective of this research was to evaluate the project of creating a standard test to measure the ability of higher level thinking of Pattaya schools students by evaluation model conduct and meta evaluation approach. The specific objective were to the reaction evaluation, learning evaluation, behavior evaluation and the result evaluation by using the integration model of Kirkpatrick and Meta evaluation approach.

The sample of this research was 114 trainees who participate in Pattaya Project on developing standard test of the Ability of Higher Thinking. The research instrument was the project evaluation forms. The statistics for data analysis were percentage, the average, standard deviation, t-test, One-way ANOVA and Pearson's Correlation Coefficient.

The result of the study concerning the reaction evaluation was that the trainees were satisfied with the facilities of the training and they suggested that the training period was too long. Whereas the result of the research in learning evaluation, the trainees were satisfied, that they have gained could gain more knowledge and more understanding on the method of creating the test basing on the Basic Education Core Curriculum. The result of the study concerning the behavior evaluation, the trainees reported that they have applied their knowledge from the training into their teaching and it was possible that the standard tests could validly measure the real students' ability. Finally, the result of the research in the study evaluation, the trainees were satisfied and they believed that the knowledge, the technique and process of test development were applicable.

# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ช
สารบัญภาพ.....	ฉ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
วัตถุประสงค์ของการวิจัย.....	3
ประโยชน์ที่ได้รับจากการวิจัย.....	4
กรอบแนวคิดในการวิจัย.....	4
ขอบเขตของการวิจัย.....	6
นิยามศัพท์เฉพาะ.....	7
2 เอกสาร และงานวิจัยที่เกี่ยวข้อง.....	9
แนวคิดและทฤษฎีที่เกี่ยวข้องกับการประเมิน โครงการ.....	10
การประเมินโครงการตามรูปแบบของเคิร์กแพตทริก.....	22
การประเมินอภิมาน.....	27
แบบทดสอบ.....	43
งานวิจัยที่เกี่ยวข้อง.....	50
3 วิธีดำเนินการวิจัย.....	54
ประชากร และกลุ่มตัวอย่าง.....	54
เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล.....	55
วิธีการสร้างเครื่องมือ.....	56
การเก็บรวบรวมข้อมูล.....	57
การวิเคราะห์ข้อมูล.....	58
สถิติที่ใช้ในการวิจัย.....	58

## สารบัญ (ต่อ)

บทที่	หน้า
4 ผลการวิเคราะห์ข้อมูล.....	61
ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมิน.....	62
ตอนที่ 2 ผลการวิเคราะห์ความพึงพอใจของผู้เข้าร่วม โครงการสร้าง ข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียน ในสังกัดเมืองพัทยา .....	66
ตอนที่ 3 การวิเคราะห์เปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบ มาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียน ในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพททริกและ การประเมินอภิมาน.....	76
ตอนที่ 4 การวิเคราะห์ความสัมพันธ์.....	90
5 สรุป อภิปรายผล และข้อเสนอแนะ.....	91
สรุปผลการวิจัย.....	91
อภิปรายผล.....	94
ข้อเสนอแนะ.....	98
บรรณานุกรม.....	100
ภาคผนวก.....	104
ภาคผนวก ก.....	105
ภาคผนวก ข.....	113
ภาคผนวก ค.....	120
ภาคผนวก ง.....	125
ประวัติย่อของผู้วิจัย.....	130

## สารบัญตาราง

ตารางที่	หน้า
1	ความสัมพันธ์ของเกณฑ์การประเมินอภิमानและขั้นตอนการประเมิน..... 38
2	ความสัมพันธ์ของเกณฑ์การประเมินอภิमानตามแนวคิดของ Stufflebeam et al. และแนวคิดของ Scriven..... 41
3	กรอบแนวคิดในการประเมิน..... 60
4	จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยาที่ตอบแบบประเมินจำแนกตาม เพศ อายุ และประสบการณ์การสอน..... 62
5	จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐาน วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยาที่ตอบแบบประเมินจำแนกตามตำแหน่งวิทยฐานะ และระดับชั้นที่สอน..... 63
6	จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยาที่ตอบแบบประเมินจำแนกตามกลุ่มสาระการเรียนรู้..... 64
7	จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยาที่ตอบแบบประเมินจำแนกตามวุฒิสูงสุด และจำแนกตามสาขา..... 65
8	ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยาด้านปฏิกิริยาตอบสนอง..... 66
9	สรุปความคิดเห็นในด้านความเหมาะสมของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา..... 67
10	ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยาด้านการเรียนรู้..... 68
11	สรุปความคิดเห็นในด้านเนื้อหาของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา..... 69



## สารบัญตาราง (ต่อ)

ตารางที่	หน้า
12 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โครงการสร้างข้อสอบมาตรฐาน วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ด้านพฤติกรรม.....	70
13 สรุปความคิดเห็นในด้านความความเป็นไปได้หลังเข้ารับการอบรม โครงการสร้าง ข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียน ในสังกัดเมืองพัทยา .....	71
14 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โครงการสร้างข้อสอบมาตรฐาน วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ด้านผลลัพธ์ที่เกิดต่อองค์กร.....	72
15 สรุปความคิดเห็นในด้านประโยชน์ของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา.....	73
16 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โดยภาพรวมของโครงการสร้าง ข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียน ในสังกัดเมืองพัทยา รายด้าน.....	75
17 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามเพศ.....	76
18 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยารายด้านและ โดยรวมของผู้เข้าอบรมที่มีเพศต่างกัน โดยการทดสอบค่าที่.....	77
19 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามช่วงอายุ.....	78
20 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามประสบการณ์สอน.....	79

## สารบัญตาราง (ต่อ)

ตารางที่	หน้า
21 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามตำแหน่ง/ วิทยฐานะ.....	80
22 การเปรียบเทียบค่าเฉลี่ย ความพึงพอใจของผู้เข้าอบรมต่อ โครงการสร้างข้อสอบ มาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนใน สังกัดเมืองพัทยา รายด้านและ โดยรวมของตำแหน่ง/ วิทยฐานะต่างกัน โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว .....	81
23 การเปรียบเทียบความพึงพอใจต่อ โครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามระดับชั้นที่สอน.....	82
24 การเปรียบเทียบค่าเฉลี่ย ความพึงพอใจของผู้เข้าอบรมต่อ โครงการสร้างข้อสอบ มาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัด เมืองพัทยา รายด้านและ โดยรวมของระดับชั้นที่สอนต่างกัน โดยการวิเคราะห์ ความแปรปรวนแบบทางเดียว .....	83
25 การเปรียบเทียบความพึงพอใจต่อ โครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามกลุ่มสาระการเรียนรู้.....	84
26 การเปรียบเทียบค่าเฉลี่ย ความพึงพอใจของผู้เข้าอบรมต่อ โครงการสร้าง ข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียน ในสังกัดเมืองพัทยา รายด้านและ โดยรวมของกลุ่มสาระการเรียนรู้ต่างกัน โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว.....	85
27 การเปรียบเทียบความพึงพอใจต่อ โครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามวุฒิการศึกษาสูงสุด.....	86
28 การเปรียบเทียบความพึงพอใจต่อ โครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา รายด้าน และ โดยรวมของผู้เข้าอบรมที่มีวุฒิการศึกษาต่างกัน โดยการทดสอบค่าที่.....	87

## สารบัญตาราง (ต่อ)

ตารางที่		หน้า
29	การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา จำแนกวุฒิตามสาขา.....	88
30	การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา รายด้าน และโดยรวมของผู้เข้าอบรมจำแนกวุฒิตามสาขา โดยการทดสอบค่าที่.....	89
31	ผลการวิเคราะห์ความสัมพันธ์ด้านปฏิริยาตอบสนอง (1), ด้านการเรียนรู้ (2), ด้านพฤติกรรม (3), ด้านผลลัพธ์ที่เกิดกับองค์กร (4) และด้านรวมกับอายุ และประสบการณ์ของผู้เข้าร่วมโครงการโดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์ แบบเพียร์สัน.....	90

## สารบัญภาพ

ภาพที่	หน้า
1 กรอบแนวคิดในการวิจัย.....	5
2 นิยามของการประเมินเพื่อการตัดสินใจ.....	11
3 สมการการประเมินอภิมานในเชิงความก้าวหน้า.....	30

# บทที่ 1

## บทนำ

### ความเป็นมาและความสำคัญของปัญหา

ประเทศไทยในปัจจุบันมีความเจริญก้าวหน้าทางเทคโนโลยีและการสื่อสาร ก่อให้เกิดการเปลี่ยนแปลงหลายด้านทั้งทางเศรษฐกิจ สังคม การเมือง และสิ่งแวดล้อมอย่างรวดเร็ว และให้ความสำคัญกับการพัฒนาประเทศโดยมุ่งเน้นทางด้านเศรษฐกิจ วัตถุประสงค์ของ ส่งผลให้คนไทยส่วนใหญ่ได้รับความสะดวกสบายจนมีค่านิยมทางด้านวัตถุสูงขึ้น ทำให้วิถีและการดำเนินชีวิตมีการเปลี่ยนแปลงจากความเรียบง่าย เอื้อเพื่อเพื่อแม่ ไปสู่ความสับสนวุ่นวาย มีการแข่งขันอยู่ตลอดเวลา ผู้คนมุ่งเน้นแต่ประโยชน์ส่วนตนมากกว่าส่วนรวม ขาดการช่วยเหลือซึ่งกันและกัน จึงกล่าวได้ว่าเจริญแต่เพียงวัตถุ แต่ขาดความเจริญทางด้านจิตใจ

การจัดการศึกษาจำเป็นต้องมีจุดมุ่งหมายที่ชัดเจน เพื่อช่วยกำหนดคุณลักษณะของผู้สำเร็จการศึกษาที่พึงประสงค์หลังจากได้รับการเรียนการสอน ช่วยให้ทิศทางของการจัดประสบการณ์การเรียนรู้หรือกิจกรรมการเรียนการสอนที่เหมาะสม การวัดและประเมินผลเข้ามามีบทบาทสำคัญในการตัดสินผลการเรียนรู้ของผู้เรียนว่าบรรลุผลสำเร็จตามจุดมุ่งหมายที่ต้องการหรือไม่ มากน้อยเพียงไร รวมทั้งช่วยให้สารสนเทศอันเป็นประโยชน์ต่อการปรับปรุงและพัฒนาการเรียนการสอน (ขวัญภา อุดมกานต์, 2553, หน้า 1)

ในปัจจุบันการประเมินและติดตามผลโครงการ เป็นสิ่งที่สำคัญในขั้นตอนหนึ่งของการดำเนินโครงการ เป็นการเก็บรวบรวมข้อมูลเพื่อนำมาใช้ในการตัดสินโครงการ นำไปสู่การแก้ปัญหา การวางแผนดำเนินการ การขยายผล และการพัฒนาแก้ไขปรับปรุงโครงการ นอกจากนี้ การประเมินและติดตามผลโครงการยังสะท้อนถึงระดับความสำเร็จของแผนงานหรือนโยบายการบริหารด้วย โดยให้ข้อมูลว่า การดำเนินโครงการบรรลุเป้าหมายวัตถุประสงค์เพียงใด คู่กับทุนและแรงงานที่ลงทุนไปหรือไม่ มีจุดแข็ง จุดอ่อน ผลสำเร็จ ผลกระทบต่อโครงการเป็นอย่างไร การประเมินและติดตามผลโครงการจะต้องมีการดำเนินการกันอย่างเป็นระบบและมีความเป็นปรนัย (จุฑารัตน์ ทนน์ไชย, 2552, หน้า 1)

แบบจำลองการประเมินและติดตามผลการฝึกอบรมของเคิร์กแพททริก เป็นอีกหนึ่งแบบจำลองที่นิยมใช้ในการประเมินและติดตามผลโครงการฝึกอบรมโดยทั่วไป โดยมีการวัดผลในระหว่างและสิ้นสุดการฝึกอบรมและมีการติดตามผลต่อเนื่องในเรื่องของการเปลี่ยนแปลงในเชิงการเรียนรู้และพฤติกรรมที่เกิดขึ้นของผู้เข้าร่วมการฝึกอบรม โดยจะประเมินเพื่อเป็นข้อมูลให้แก่

ผู้บริหารหรือผู้รับผิดชอบโครงการ ผู้เข้าร่วมฝึกอบรม และวิทยากรให้ทราบถึงผลที่ได้รับจากการฝึกอบรมเพื่อนำไปสู่การปรับปรุงเปลี่ยนแปลงให้โครงการฝึกอบรมมีประสิทธิภาพ และเหมาะสมกับผู้เข้าร่วมโครงการในแต่ละครั้ง แนวคิดของเคิร์กแพททริกมีการประเมิน 4 ขั้นตอน โดยกำหนดเป็น Levels of training evaluation ดังนี้ ขั้นปฏิกิริยาตอบสนอง (Reaction evaluation) ขั้นการเรียนรู้ (Learning evaluation) ขั้นพฤติกรรม (Behavior evaluation) และขั้นผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) โดยรูปแบบนี้เชื่อว่าการฝึกอบรมจะทำให้เกิดการเปลี่ยนแปลงในบุคคลและหน่วยงาน รูปแบบนี้มีจุดเด่น คือ โดยทั่วไปคนนิยมใช้เป็นรูปแบบในการประเมินและติดตามผลโครงการฝึกอบรมเนื่องจากรูปแบบจะเน้นวัดในขั้นตอนสุดท้าย จึงเป็นกระบวนการที่ง่ายและสะดวกต่อการประเมินตามผล และแบบจำลองนี้จะเป็นการประเมินที่มุ่งสัมฤทธิ์ผลของการฝึกอบรมในหลาย ๆ เรื่องที่มีความสัมพันธ์กันในลักษณะขึ้นบันได ทำให้เห็นการเปลี่ยนแปลงที่เกิดขึ้นกับตัวบุคคลและหน่วยงาน ส่วนข้อจำกัด คือ แบบจำลองนี้มีการประเมิน 4 ขั้นตอน แต่โดยทั่วไปผู้ประเมินจะเลือกใช้เพียง 2 ขั้นตอนแรก คือ ขั้นปฏิกิริยาตอบสนอง (Reaction evaluation) และขั้นการเรียนรู้ (Learning evaluation) จะไม่ประเมินครบทุกขั้นตอน เนื่องจากขั้นที่ 3 และ 4 คือ ขั้นพฤติกรรม (Behavior evaluation) และขั้นผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) เป็นการติดตามผลว่าผู้เข้าร่วมการอบรมจะนำความรู้ที่ได้รับไปใช้หรือไม่ แล้วเกิดการเปลี่ยนแปลงต่อบุคคลและองค์กรอย่างไร โดยจะมีการเว้นระยะเวลาไว้ เช่น 3 เดือน เป็นต้น หลังจากนั้นมีการวัดผลที่เกิดจากการเปลี่ยนแปลง

การประเมินอภิมาน (Meta evaluation) เป็นวิธีการประเมินงานประเมินที่แพร่หลายในต่างประเทศมาเป็นเวลานาน สำหรับในประเทศไทยนั้นยังเป็นที่รู้จักไม่มากนักจะรู้จักกันในวงจำกัดเพียงเฉพาะกลุ่มอาจารย์และนักวิชาการด้านการวัดและประเมินผลทางการศึกษาเท่านั้น และเริ่มนำมาประยุกต์ใช้เพื่อตรวจสอบคุณภาพของงานประเมินทางการศึกษา (เนติ เกลียวเรศ, 2556, หน้า 1)

การประเมินอภิมาน ด้วยแบบตรวจสอบสำหรับการประเมินอภิมานรายงานประเมินโครงการ (Program evaluation meta-evaluation checklist) ที่พัฒนาขึ้นโดยสตัฟเฟิลบีม (Stufflebeam) ได้ดำเนินการจัดทำมาตรฐานสำหรับการตัดสินคุณภาพการประเมินทางการศึกษา ซึ่งแบ่งออกเป็น 4 มาตรฐาน คือ มาตรฐานด้านการใช้ประโยชน์ (Utility standards) มาตรฐานด้านความเป็นไปได้ (Feasibility standards) มาตรฐานด้านความเหมาะสม (Propriety standards) และมาตรฐานด้านความถูกต้อง (Accuracy standards) จำนวน 30 เกณฑ์ มีความครอบคลุมภาระงานของการประเมินในทุกขั้นตอนเริ่มตั้งแต่การตัดสินใจที่จะประเมินการกำหนดเป้าหมายของการประเมินการเป็นการเป็นที่ยอมรับได้ทางการเมือง การทำสัญญา การจัดทีมผู้ประเมิน

การบริหารจัดการ/ ดำเนินการประเมิน การรวบรวมข้อมูล การวิเคราะห์ข้อมูล การรายงาน  
 ข้อค้นพบและการนำผลการประเมินไปใช้ ดังนั้นผู้วิจัยจึงเลือกใช้เกณฑ์การประเมิน ดังกล่าวมาใช้ในการประเมินอภิमानงานวิจัยที่เกี่ยวกับการประเมินต่อไป

เมื่อพิจารณารูปแบบการประเมินทั้งสองรูปแบบ พบว่า มีส่วนที่สอดคล้องและ  
 คล้ายคลึงกัน จึงนำมาบูรณาการร่วมกันเพื่อใช้ในการประเมิน โครงการสร้างข้อสอบมาตรฐานวัด  
 ความรู้ ความสามารถในการคิดขั้นสูง ซึ่งมีอยู่ 4 ด้าน ได้แก่ การประเมินด้านปฏิบัติการความพึงพอใจ  
 ร่วมกับความเหมาะสมของโครงการ การประเมินด้านการเรียนรู้ทักษะการปฏิบัติร่วมกับ  
 ความถูกต้องของเนื้อหาในการอบรม การประเมินด้านพฤติกรรมการนำความรู้ที่ได้ไปปฏิบัติจริง  
 ร่วมกับความเข้าใจได้หลังจากอบรมเสร็จแล้ว และการประเมินด้านผลลัพธ์ที่พึงประสงค์ของ  
 โครงการร่วมกับประโยชน์ในการอบรม

จากความสำคัญดังกล่าวจะเห็นว่าทักษะการสร้างข้อสอบของครูเป็นสิ่งสำคัญ  
 ที่มีส่วนช่วยในการพัฒนาผลสัมฤทธิ์ทางการเรียนของนักเรียนทั้งการทดสอบในระดับชั้นเรียน  
 การทดสอบในระดับท้องถิ่น หรือแม้กระทั่งการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-NET)  
 สำหรับการวิจัยในครั้งนี้ ผู้วิจัยได้ประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถ  
 ในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมิน  
 ของเคิร์กแพตทริกและการประเมินอภิमान เพื่อนำข้อมูลที่ได้จากการประเมินมาใช้ปรับปรุง  
 การวางแผนพัฒนา ปรับปรุง โครงการ และการดำเนินงานของการจัดทำแบบทดสอบของ โรงเรียน  
 ในสังกัดเมืองพัทยา อันจะเป็นประโยชน์ต่อการพัฒนาปรับปรุงแก้ไข และเป็นประเมินเพื่อช่วย  
 ในการตัดสินใจเกี่ยวกับการวางแผนดำเนินงาน โครงการในปีการศึกษาต่อไป

### วัตถุประสงค์ของการวิจัย

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้  
 ความสามารถในการคิดขั้นสูงของนักเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมิน  
 ของเคิร์กแพตทริกและการประเมินอภิमानใน 4 ด้าน ดังนี้

1. ปฏิกริยาตอบสนอง (Reaction evaluation) ตามรูปแบบการประเมินของ  
 เคิร์กแพตทริกและการประเมินอภิमान
2. การเรียนรู้ (Learning evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริก  
 และการประเมินอภิमान
3. พฤติกรรม (Behavior evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริก  
 และการประเมินอภิमान

4. ผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) ตามรูปแบบการประเมินของ  
 เครื่องแพดทริกและการประเมินอภิमान

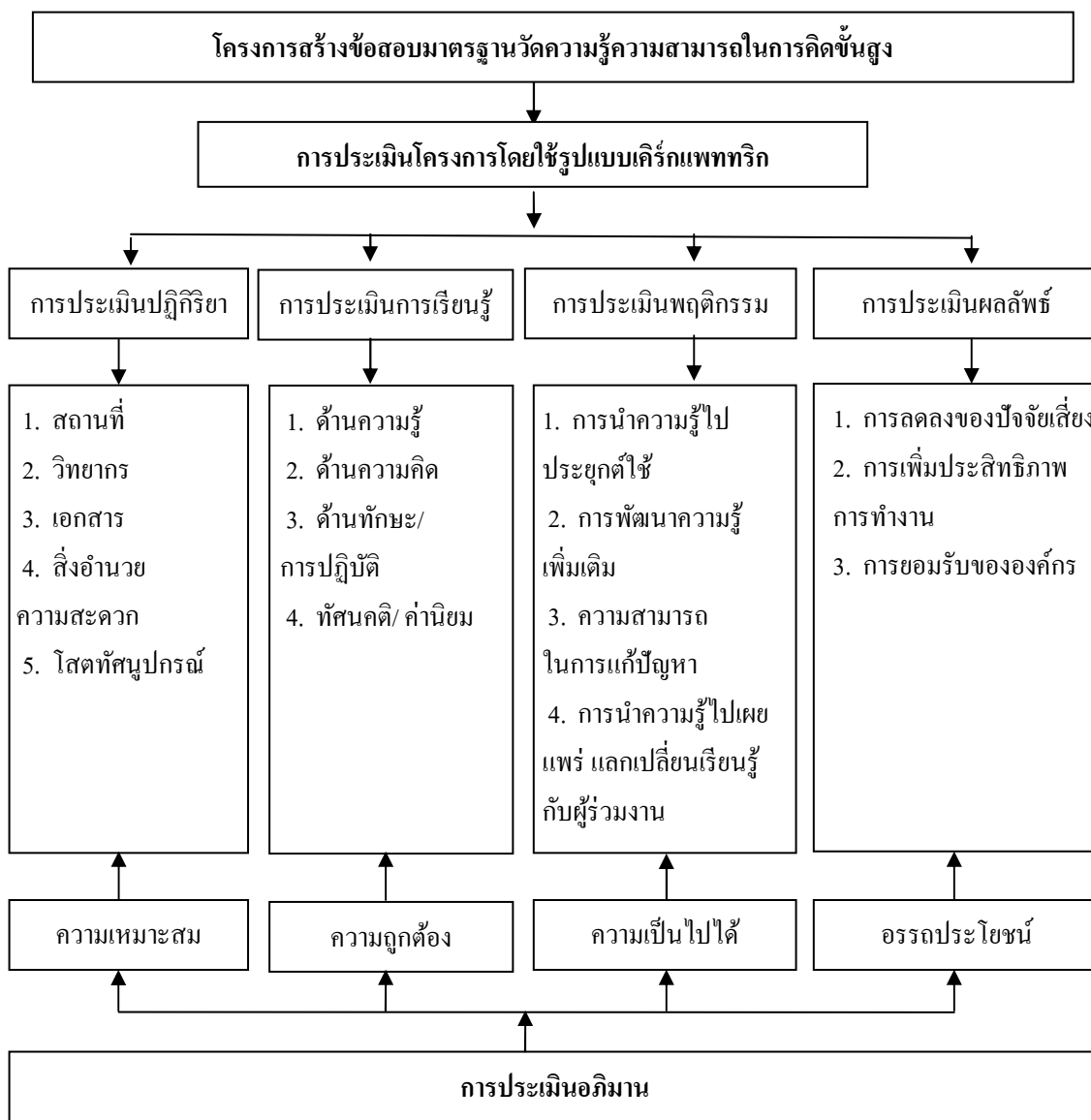
### ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อนำข้อมูลที่ได้จากการประเมินมาใช้ปรับปรุง การวางแผนพัฒนา ปรับปรุง  
 โครงการ และการดำเนินงานของการจัดทำแบบทดสอบของ โรงเรียนในสังกัดเมืองพัทยา
2. ทำให้ทราบถึงข้อเสนอแนะโครงการ อันจะนำไปสู่การบริหารจัดการ โครงการสร้าง  
 ข้อสอบมาตรฐานเพื่อวัดความรู้ ความสามารถในการคิดขั้นสูงให้มีประสิทธิภาพยิ่งขึ้นและพัฒนา  
 วางแผนดำเนินงานโครงการในปีการศึกษาต่อไป

### กรอบแนวคิดในการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยได้บูรณาการรูปแบบการประเมินของเครื่องแพดทริกที่เน้น  
 การประเมินหลังโครงการ 4 ระดับ คือ การประเมินปฏิกิริยาตอบสนอง (Reaction evaluation)  
 การประเมินการเรียนรู้ (Learning evaluation) การประเมินพฤติกรรม (Behavior evaluation)  
 และการประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) กับการประเมินอภิमान ซึ่งแบ่งได้เป็น  
 4 มิติ คือ รรถประโยชน์ (Utility) ความเป็นไปได้ (Feasibility) ความเหมาะสม (Propriety)  
 และความถูกต้อง (Accuracy)





ภาพที่ 1 กรอบแนวคิดในการวิจัย

จากภาพสามารถอธิบายได้ว่า ในการวิจัยครั้งนี้ผู้วิจัยได้ประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง โดยบูรณาการรูปแบบการประเมินของ เคิร์กแพททริกและการประเมินอภิมาน ซึ่งมีอยู่ 4 ด้าน ได้แก่ การประเมินด้านปฏิบัติการ ความพึงพอใจร่วมกับความเหมาะสมของโครงการ การประเมินด้านการเรียนรู้ทักษะการปฏิบัติ ร่วมกับความต้องการของเนื้อหาในการอบรม การประเมินด้านพฤติกรรมการนำความรู้ที่ได้ไปปฏิบัติจริงร่วมกับความเป็นไปได้หลังจากอบรมเสร็จแล้ว และการประเมินด้านผลลัพธ์ ที่พึงประสงค์ของโครงการร่วมกับประโยชน์ในการอบรม

## ขอบเขตของการวิจัย

### ขอบเขตด้านเนื้อหา

การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของเมืองพัทยาโดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน

#### 1. รูปแบบการประเมินของเคิร์กแพตทริก 4 ระดับ ดังนี้

1.1 การประเมินปฏิกิริยาตอบสนอง (Reaction evaluation) เพื่อให้ทราบว่าผู้เข้าอบรมมีความพึงพอใจต่อโครงการหรือไม่ มากน้อยเพียงใด

1.2 การประเมินการเรียนรู้ (Learning evaluation) เพื่อให้ทราบว่าผู้เข้าอบรมได้เรียนรู้อะไรจากโครงการฝึกอบรมบ้าง

1.3 การประเมินพฤติกรรม (Behavior evaluation) เพื่อให้ทราบว่าผู้เข้าอบรมเปลี่ยนแปลงพฤติกรรมอันเนื่องมาจากสิ่งได้เรียนรู้มาหรือไม่

1.4 การประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) เพื่อให้รู้ว่าการเปลี่ยนแปลงพฤติกรรมมีผลกระทบในทางบวกต่อองค์กรหรือไม่

#### 2. การประเมินอภิมานตามแนวคิดของ Stufflebeam et al.

คณะกรรมการกำหนดมาตรฐานการประเมินทางการศึกษา (The joint committee on standards for educational evaluation) ซึ่งมี Stufflebeam เป็นผู้อำนวยการได้ดำเนินการจัดทำมาตรฐานสำหรับการตัดสินคุณภาพการประเมินทางการศึกษา ซึ่งแบ่งออกเป็น 4 มาตรฐาน คือ

2.1 มาตรฐานด้านการใช้ประโยชน์ (Utility standards) เป็นการประเมินด้านกระบวนการที่ก่อให้เกิดสารสนเทศตรงกับความต้องการของผู้ใช้ผลการประเมินเป็นสำคัญ

2.2 มาตรฐานด้านความเป็นไปได้ (Feasibility standards) เป็นมาตรฐานเน้นความเป็นจริงของการประเมิน สามารถนำไปใช้ปฏิบัติได้จริง เป็นที่ยอมรับของผู้ถูกประเมิน ผู้ได้รับผลกระทบจากการประเมิน ตลอดจนผู้ที่เกี่ยวข้องทุกฝ่าย สามารถนำไปปรับปรุงการปฏิบัติงาน

2.3 มาตรฐานด้านความเหมาะสม (Propriety standards) มุ่งเน้นเกี่ยวกับเรื่อง กฎระเบียบ ศีลธรรม จรรยาบรรณที่เกี่ยวข้องกับการประเมิน

2.4 มาตรฐานด้านความถูกต้อง (Accuracy standards) เกี่ยวข้องกับการใช้เครื่องมือที่มีคุณภาพ ตลอดจนวิธีเก็บข้อมูลตามหลักวิชาการ

### ขอบเขตด้านประชากร

ประชากรที่ใช้ในการประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของเมืองพัทยาโดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและ

การประเมินอภิมาน จำนวน 609 คน กลุ่มตัวอย่างที่ใช้เป็นผู้เข้าอบรมในโครงการ จำนวน 114 คน ประกอบด้วย หัวหน้าฝ่ายวิชาการจำนวน 1 คน นักวิชาการศึกษา/ เจ้าหน้าที่โครงการ จำนวน 3 คน ผู้อำนวยการสถานศึกษา/ รองผู้อำนวยการสถานศึกษา (รับผิดชอบงานวิชาการ) จำนวน 10 คน ผู้แทนครูวัดผลทุกสถานศึกษาจำนวน 11 คน ครูผู้สอนกลุ่มสาระการเรียนรู้ทั้ง 8 กลุ่มสาระการเรียนรู้ จำนวน 89 คน

### นิยามศัพท์เฉพาะ

การประเมินปฏิกิริยาตอบสนอง (Reaction evaluation) หมายถึง การประเมินของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง ว่ามีความรู้ดีอย่างไรต่อการฝึกอบรมในด้านความเหมาะสมของสถานที่ วิทยากร เอกสารประกอบ สิ่งอำนวยความสะดวก และด้านความเหมาะสมโสตทัศนูปกรณ์

การประเมินการเรียนรู้ (Learning evaluation) หมายถึง การประเมินของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง ในด้านความรู้ ด้านความคิด ด้านทักษะ/ การปฏิบัติ ทักษะคิด/ ค่านิยม ที่ทำให้รู้ว่าผู้เข้ารับการฝึกอบรมได้รับความรู้และทักษะอะไรบ้าง มีเจตคติอะไรบ้างที่เปลี่ยนแปลงไปจากเดิมทั้งก่อนและหลัง

การประเมินพฤติกรรม (Behavior evaluation) หมายถึง การประเมินของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง ในด้านการนำความรู้ไปประยุกต์ใช้ การพัฒนาความรู้เพิ่มเติม ความสามารถในการแก้ปัญหา การนำความรู้ไปเผยแพร่ แลกเปลี่ยนเรียนรู้กับผู้ร่วมงาน เพื่อตรวจสอบว่าผู้เข้ารับการฝึกอบรมได้มีการเปลี่ยนแปลงพฤติกรรมในการทำงานไปในทิศทางที่พึงประสงค์หรือไม่

การประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) หมายถึง การประเมินของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง ในด้านการลดลงของปัจจัยเสี่ยง การเพิ่มประสิทธิภาพการทำงาน การยอมรับขององค์กร

การประเมินอภิมาน (Meta-evaluation) หมายถึง กระบวนการตรวจสอบคุณภาพของประเมิน โดยอาศัยเกณฑ์ในการประเมินที่มีความครอบคลุมกระบวนการประเมินเพื่อทำให้ทราบถึงคุณภาพของงานประเมินและแนวทางในการปรับปรุงการประเมินสามารถแบ่งออกได้เป็น 2 ประเภท คือ การประเมินอภิมานเชิงความก้าวหน้า (Formative meta-evaluation) และการประเมินอภิมานเชิงสรุปรวม (Summative meta-evaluation)

ผู้มีส่วนได้ส่วนเสีย หมายถึง ผู้บริหาร คณะครู คณะเข้ารับการฝึกอบรม

การบูรณาการรูปแบบการประเมิน หมายถึง การบูรณาการรูปแบบการประเมินของ  
เครื่องเพศทริกกับการประเมินอภิมาน แบ่งได้ ดังนี้ การประเมินด้านปฏิริยาความพึงพอใจร่วมกับ  
ความเหมาะสมของโครงการ การประเมินด้านการเรียนรู้ทักษะการปฏิบัติร่วมกับความถูกต้อง  
ของเนื้อหาในการอบรม การประเมินด้านพฤติกรรมการนำความรู้ที่ได้ไปปฏิบัติจริงร่วมกับ  
ความเป็นไปได้หลังจากอบรมเสร็จแล้ว และการประเมินด้านผลลัพธ์ที่พึงประสงค์ของโครงการ  
ร่วมกับประโยชน์ในการอบรม

## บทที่ 2

### เอกสาร และงานวิจัยที่เกี่ยวข้อง

การศึกษาวิจัยเรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน” ผู้วิจัยได้ศึกษาทบทวนเอกสาร รวบรวมแนวคิด ทฤษฎี และผลงานวิจัยที่เกี่ยวข้อง เพื่อเป็นแนวทางในการศึกษา ดังต่อไปนี้

1. แนวคิดและทฤษฎีที่เกี่ยวข้องกับการประเมินโครงการ
  - 1.1 ความหมายของการประเมินโครงการ
  - 1.2 ความมุ่งหมายและความสำคัญของการประเมินโครงการ
  - 1.3 ประเภทของการประเมินโครงการ
  - 1.4 รูปแบบของการประเมินโครงการ
  - 1.5 ประโยชน์ของการประเมินโครงการ
2. การประเมินโครงการตามรูปแบบของเคิร์กแพตทริก
  - 2.1 การประเมินปฏิกิริยาตอบสนอง (Reaction evaluation)
  - 2.2 การประเมินการเรียนรู้ (Learnig evaluation)
  - 2.3 การประเมินพฤติกรรม (Behavior evaluation)
  - 2.4 การประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation)
3. การประเมินอภิมาน
  - 3.1 ความหมายของการประเมินอภิมาน
  - 3.2 จุดมุ่งหมายของการประเมินอภิมาน
  - 3.3 ประเภทของการประเมินอภิมาน
  - 3.4 หลักการประเมินอภิมาน
  - 3.5 ขั้นตอนของการประเมินอภิมาน
  - 3.6 เกณฑ์การประเมินอภิมาน
4. แบบทดสอบ
  - 4.1 ความหมายของแบบทดสอบ
  - 4.2 ประเภทของแบบทดสอบ
  - 4.3 แบบทดสอบมาตรฐาน

- 4.4 ขั้นตอนการสร้างข้อสอบมาตรฐาน
- 4.5 องค์ประกอบของแบบทดสอบมาตรฐาน
5. งานวิจัยที่เกี่ยวข้อง

## แนวคิดและทฤษฎีที่เกี่ยวข้องกับการประเมินโครงการ

### ความหมายของการประเมินโครงการ

การประเมินโครงการ (Project evaluation) เป็นคำที่มาจาก การประเมินผล (Evaluation) กับคำว่าโครงการ (Project) นักวิชาการได้ให้ความหมายไว้ ดังนี้

สมหวัง พิธิยานุวัฒน์ (2549, หน้า 6) กล่าวว่า เป็นการตัดสินคุณค่าของสิ่งใดสิ่งหนึ่ง ซึ่งถือเป็นนิยามพื้นฐานในทางการจัดการ นิยมนิยามการประเมินค่าหรือการประเมินผลว่าเป็นกระบวนการที่ก่อให้เกิดสารสนเทศเพื่อการตัดสินใจ การตัดสินใจเลือกทางเลือกโดยอาศัยสารสนเทศที่ถูกต้องเหมาะสม เมื่อผ่านการสังเคราะห์ให้เป็นองค์ความรู้ จะทำให้เกิดปัญญาได้

พิชญ์ ฟองสี (2550, หน้า 4) ได้ให้ความหมายของการประเมินไว้ว่า เป็นกระบวนการตัดสินคุณค่าของสิ่งหนึ่งสิ่งใดโดยนำสารสนเทศหรือผลจากการวัดมาเปรียบเทียบกับเกณฑ์ที่กำหนดไว้

สมคิด พรหมจ้อย (2550, หน้า 37) ให้ความหมายที่สะท้อนถึงทั้งวัตถุประสงค์และประเภทของการประเมินโครงการว่าเป็นการตรวจสอบความก้าวหน้าของโครงการหรือแผนงาน ตลอดจนการพิจารณาผลสัมฤทธิ์ของโครงการหรือแผนงานนั้น ๆ ว่ามีมากน้อยเพียงใด การประเมินผลเป็นกระบวนการบ่งชี้ถึงคุณค่าของโครงการ กล่าว คือ โครงการที่ได้ดำเนินการไปแล้ว ได้ผลตามวัตถุประสงค์ของโครงการหรือไม่ เพียงใด สามารถทำได้ทั้งการประเมินก่อนเริ่มโครงการ การประเมินขณะที่โครงการกำลังดำเนินการอยู่ และการประเมินผลโครงการหลังจากการดำเนินงานได้สิ้นสุดแล้ว

เชาว์ อินโย (2553, หน้า 4) ได้ให้ความหมายของการประเมิน หมายถึง กระบวนการพิจารณาตัดสินคุณค่าของสิ่งใดสิ่งหนึ่งว่า มีความเหมาะสมหรือไม่เพียงใด โดยนำสารสนเทศหรือผลจากการวัดมาเปรียบเทียบกับเกณฑ์ที่กำหนดเพื่อช่วยในการตัดสินใจ ตีค่าผลการดำเนินการนั้น ๆ ว่าบรรลุวัตถุประสงค์หรือไม่ ใช้เป็นส่วนหนึ่งของกระบวนการจัดการ ส่วนคำว่าโครงการ หมายถึง ส่วนย่อยส่วนหนึ่งของแผนงาน ซึ่งประกอบด้วย ชุดของกิจกรรมที่จัดขึ้นอย่างมีระบบ มีการกำหนดทรัพยากรในการดำเนินงาน ระยะเวลาดำเนินงานไว้อย่างชัดเจน โดยออกแบบมาเพื่อให้บรรลุเป้าหมายตามต้องการ

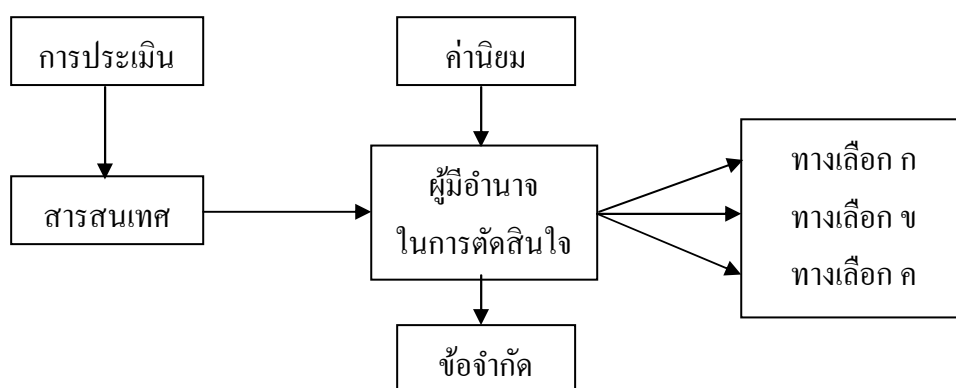
กัจจตุ สุตโต (2553, หน้า 11) กล่าวว่า การประเมินเป็นกระบวนการรวบรวมสารสนเทศ โดยมีการรวบรวมข้อมูลและวิธีการศึกษาอย่างเป็นระบบระเบียบเพื่อเปรียบเทียบการปฏิบัติงาน กับเป้าหมายที่กำหนดไว้ว่าการดำเนินงานบรรลุวัตถุประสงค์หรือไม่โดยนำผลที่ได้มาพิจารณา วิเคราะห์เปรียบเทียบกับแผนที่กำหนดไว้ ผลที่ได้จากการประเมินนี้จะสามารถชี้ให้เห็นถึงระดับ ความสำเร็จ ความผิดพลาด ปัญหา อุปสรรค ผลกระทบต่อโครงการในอนาคตที่จะนำไปปรับปรุง การดำเนินงานตามแผนงานโครงการให้มีประสิทธิภาพต่อไป

ขวัญภา อุณหกานต์ (2553, หน้า 10-11) การประเมินโครงการ หมายถึง กระบวนการ ในการเก็บรวบรวมข้อมูลและวิเคราะห์ข้อมูลของการดำเนินโครงการและพิจารณาปัจจัยให้ทราบถึง จุดเด่นหรือจุดด้อยของโครงการอย่างมีระบบแล้วตัดสินใจว่าจะปรับปรุงแก้ไขโครงการนั้นเพื่อ การดำเนินงานต่อไป หรือจะยุติการดำเนินงานโครงการนั้นเสีย หรือนำข้อมูล ปัญหา อุปสรรคที่พบ ในการประเมินโครงการนำไปใช้เป็นแนวทางดำเนินโครงการที่มีลักษณะที่ใกล้เคียงกันต่อไป

สมหวัง พิธิยานุวัฒน์ (2553, หน้า 20-21) ให้ความหมายว่า การประเมิน หมายถึง กระบวนการใช้ดุลย-พินิจ (Judgment) หรือค่านิยมและข้อจำกัดต่าง ๆ ในการพิจารณาตัดสินคุณค่า ของสิ่งใดสิ่งหนึ่ง โดยการเปรียบเทียบผลที่วัดได้กับเกณฑ์ที่กำหนดไว้ ซึ่งสรุปเป็นสมการได้ ดังนี้

$$\text{การประเมิน} = \text{การวัด} + \text{ดุลยพินิจ}$$

การประเมินจึงเป็นกระบวนการที่ก่อให้เกิดสารสนเทศ (เชิงคุณค่า) เพื่อช่วยให้ผู้มีอำนาจ ตัดสินใจได้ตัดสินใจเลือกทางเลือกของที่มีประสิทธิภาพสูงสุด ดังแสดงในภาพที่ 2



ภาพที่ 2 นิยามของการประเมินเพื่อการตัดสินใจ

พงศ์เทพ จิระโร (2554, หน้า 4) ให้ความหมายการประเมินโครงการ (Project evaluation) หมายถึง กระบวนการวัดตัวบ่งชี้คุณภาพโครงการที่กำหนดไว้ และการนำเกณฑ์ (Criterion) มาเทียบกับผลการวัด ทำให้เกิดผลการประเมินที่นำไปใช้ในการปรับปรุงและพัฒนาโครงการหรือนำไปสู่การตัดสินใจ (Judgment) ในการดำเนินโครงการ

ปยุตดา จันทรเส็ง (2557, หน้า 9) การประเมินโครงการ หมายถึง กระบวนการที่หาข้อมูลที่เป็นจริง เกี่ยวกับโครงการ โดยการเก็บรวบรวมข้อมูลเกี่ยวกับวัตถุประสงค์ การวางแผน การดำเนินการและผลกระทบที่เกิดขึ้น เพื่อเป็นการกำหนดคุณค่าหรือข้อดีของโครงการดังกล่าวว่าดีหรือไม่ อย่างไร หรือเป็นการค้นหาว่าผลสำเร็จของกิจกรรมจากการดำเนินโครงการที่กำหนดไว้ว่าสามารถ ประสบความสำเร็จตรงตามวัตถุประสงค์ที่กำหนดไว้ในแผนหรือไม่

กล่าวโดยสรุป การประเมินโครงการ หมายถึง กระบวนการเชิงระบบเพื่อการตรวจสอบหรือชี้บ่งถึงประสิทธิภาพและประสิทธิผลของโครงการซึ่งจะช่วยให้ได้ข้อมูลสารสนเทศที่เป็นประโยชน์ต่อการตัดสินใจเกี่ยวกับการบริหารจัดการ ปรับปรุงและพัฒนาโครงการ จากความหมายดังกล่าวนี้ชี้ให้เห็นลักษณะของการประเมินโครงการ ดังนี้

1. การประเมินโครงการเป็นกระบวนการที่จัดทำขึ้นอย่างเป็นระบบเพื่อให้ได้ข้อมูลสารสนเทศที่เชื่อถือได้และเป็นประโยชน์ต่อการตัดสินใจของผู้บริหารหรือผู้รับผิดชอบโครงการ
2. การประเมินโครงการเป็นกระบวนการในการจัดเตรียมสารสนเทศ เพื่อใช้ในการตัดสินใจเกี่ยวกับการบริหารจัดการ ปรับปรุงและพัฒนาโครงการ
3. จุดเน้นที่สำคัญของการประเมินโครงการ อยู่ที่การเพิ่มประสิทธิภาพ (Efficiency) และประสิทธิผล (Effectiveness) ของการดำเนินโครงการ

#### **ความมุ่งหมายและความสำคัญของการประเมินโครงการ**

มานะ ทองรักษ์ (2549, หน้า 72-73) บทบาทหน้าที่ของการศึกษาในปัจจุบัน โดยเฉพาะในโรงเรียนมีขอบข่ายขยายกว้าง และซับซ้อนมากขึ้น มีภาระมากขึ้น เพราะความเติบโตของสังคม ความต้องการของสังคมเปลี่ยนแปลงไปตามสภาพสังคม ซึ่งปัจจุบันโรงเรียนไม่ได้มีหน้าที่เฉพาะสอนนักเรียนเพียงอย่างเดียวเท่านั้น แต่ต้องเกี่ยวข้องกับชุมชน โรงเรียนจำเป็นต้องบริหารงานให้ดำเนินไปตามนโยบายของการศึกษาระดับสูง ต้องดำเนินตามแผนและประเมินผลโครงการเพื่อให้งานดำเนินไปตามจุดประสงค์ จึงทำให้การประเมินโครงการทางการศึกษามีความสำคัญดังต่อไปนี้

1. ช่วยชี้ให้เห็นว่าจุดประสงค์ของการดำเนินงานนั้นเหมาะสมและเป็นไปได้เพียงใด การดำเนินโครงการใด ๆ จึงต้องมีการกำหนดจุดประสงค์ของการดำเนินงาน การประเมินจะเป็นตัวชี้ให้เห็นถึงความจำเป็นที่ต้องดำเนินโครงการนี้และการดำเนินงานนั้นมีความเป็นไปได้เพียงใด



2. ทำให้ทราบว่าการดำเนินงานนั้นบรรลุวัตถุประสงค์หรือไม่ การประเมินโครงการนั้น นอกจากประเมินส่วนอื่น ๆ แล้ว จะต้องประเมินว่าการดำเนินงานนั้นบรรลุตามวัตถุประสงค์ หรือไม่เพื่อเป็นการตรวจสอบว่าได้ดำเนินงานไปแล้ว ได้ผลตามจุดประสงค์ที่ตั้งไว้หรือไม่ ซึ่งเป็นการประเมินผลสัมฤทธิ์ของ โครงการ

3. การกระตุ้นให้มีการเร่งรัดปรับปรุงการดำเนินงาน การประเมินจะเป็นตัวกระตุ้น ให้ผู้ดำเนินงาน มีการเร่งรัดและปรับปรุงการดำเนินงานเมื่อพบข้อบกพร่องในการดำเนินงาน

4. ช่วยให้เห็นข้อบกพร่องในการดำเนินงานแต่ละขั้นตอนซึ่งจะใช้เป็นหลัก ในการปรับปรุงการดำเนินงาน การประเมินการดำเนินงานทุกขั้นตอนจะทำให้พบข้อบกพร่อง

5. ช่วยควบคุมการดำเนินงานให้มีคุณภาพและประสิทธิภาพซึ่งจะเป็นการลด ความสูญเสียในการใช้ทรัพยากร ผู้ดำเนินงานบางคน บางครั้งมักไม่ปฏิบัติหน้าที่ให้ดีและเต็ม ความสามารถ ซึ่งการประเมินจะช่วยควบคุมการดำเนินงานให้มีคุณภาพได้

6. ช่วยให้ออกเสนอแนะผู้บริหารในด้านการดำเนินงาน ผู้บริหารต้องการทราบข้อมูล ทุกแง่มุมของผลจากการประเมิน ซึ่งผู้บริหารจะได้นำไปพิจารณาวินิจฉัยและตัดสินใจ สั่งการได้ถูกต้อง

7. ใช้เป็นแนวทางในการกำหนดวิธีการดำเนินงานที่เหมาะสมในครั้งต่อ ๆ ไป การประเมินโครงการที่ดำเนินงานในปัจจุบันจะทำให้ทราบถึงข้อดีและข้อบกพร่องต่าง ๆ ของวิธี ดำเนินงานซึ่งจะเป็นแนวทางในการแก้ไขที่จะดำเนินการประเภทนี้ต่อไป

จากการพิจารณาความหมายของการประเมิน โครงการ ซึ่งให้ เห็นว่าการประเมินโครงการ เป็นกิจกรรมสำคัญที่ขาดไม่ได้ในการบริหารจัดการ โครงการ เพราะการประเมินโครงการมี ประโยชน์อย่างยิ่งต่อการตัดสินใจของผู้บริหารและผู้รับผิดชอบโครงการ ซึ่งสรุปได้ ดังนี้

1. ช่วยให้ได้ข้อมูลสารสนเทศต่าง ๆ เกี่ยวกับโครงการที่นำไปใช้ในการตัดสินใจ เกี่ยวกับการกำหนดโครงการ การตรวจสอบความพร้อมของทรัพยากรต่าง ๆ ที่จำเป็นต้องใช้ ในการดำเนินโครงการ ตลอดจนความเป็นไปได้ในการจัดทำโครงการ

2. ช่วยให้ทราบข้อมูลสารสนเทศเกี่ยวกับความก้าวหน้า ปัญหาและอุปสรรคใน การดำเนินโครงการซึ่งนำมาใช้ในการตัดสินใจเพื่อการปรับปรุง หรือเปลี่ยนแปลงแก้ไขการดำเนิน โครงการให้เป็นไปตามทิศทางที่ต้องการ

3. ช่วยให้ได้ข้อมูลสารสนเทศเกี่ยวกับความสำเร็จและความล้มเหลวของโครงการ ที่นำมาใช้ในการตัดสินใจและวินิจฉัยว่า จะดำเนินโครงการในช่วงต่อไปอย่างไร ยกเลิกหรือปรับ ขยายการดำเนิน โครงการต่อไปอีก

4. ช่วยให้ได้ข้อมูลสารสนเทศที่บ่งบอกถึงประสิทธิภาพของโครงการว่าเป็นอย่างไร คุ่มค่ากับการลงทุนหรือไม่

5. ช่วยให้เกิดการเสริมแรง สร้างพลังใจให้กับผู้ปฏิบัติงานในโครงการเมื่อทราบผลสำเร็จ จุดเด่น หรือจุดด้อยของโครงการ โดยจะมุ่งมั่นปรับปรุง และพัฒนาการดำเนินโครงการ ให้มีคุณภาพและมาตรฐานสูงขึ้นซึ่งจะเกิดคุณค่าและประโยชน์สูงสุดต่อผู้รับบริการหรือองค์กร

กำจัด สุกโต (2553, หน้า 12) สรุปความมุ่งหมายและความสำคัญของการประเมินโครงการไว้ว่า การประเมินโครงการอย่างมีระบบสามารถกำหนดวัตถุประสงค์ของการดำเนินงาน ให้มีความชัดเจนยิ่งขึ้น กำหนดการใช้ทรัพยากรอย่างคุ้มค่าและเกิดประโยชน์สูงสุด มีส่วนสำคัญในการควบคุมคุณภาพของงาน ช่วยในการตัดสินใจการบริหารโครงการ และคงไว้ซึ่งคุณภาพและมาตรฐานในการดำเนินงาน ตลอดจนความรับผิดชอบของหน่วยงาน

เชาว์ อินโย (2553, หน้า 12) ได้อธิบายความสำคัญของการประเมินโครงการไว้ว่า การประเมินโครงการเป็นส่วนหนึ่งของการวิจัย เป็นกระบวนการที่มีระบบเพื่อค้นหาข้อเท็จจริง หรือกฎเกณฑ์ต่าง ๆ ทั้งยังเป็นกระบวนการที่มีระบบเพื่อตัดสินความสำเร็จของโครงการอีกด้วย การประเมินโครงการเป็นการดำเนินงานที่ไม่ใช้ความพยายามในการสร้างทฤษฎี หรือพัฒนาองค์ความรู้ทางสังคมศาสตร์ การประเมินโครงการที่นำมาใช้ในทางสังคมศาสตร์นั้น เป็นการเตรียมสารสนเทศเพื่อใช้ในการปรับปรุงโครงการทางสังคม เหตุผลประการสำคัญที่จำเป็นต้องประเมินโครงการก็คือ มีทางเลือกในการดำเนินโครงการได้มากมายที่จะทำให้การดำเนินงานโครงการมีประสิทธิภาพ ดังนั้น จึงจำเป็นต้องประเมินโครงการว่า ประสบความสำเร็จหรือไม่

พงศ์เทพ จิระโร (2554, หน้า 4) กล่าวเกี่ยวกับความสำคัญของการประเมินโครงการนั้น ในการประเมินโครงการแต่ละโครงการก็จะมีค่าตามบริบทของการประเมินครานั้น ๆ ซึ่งความสำคัญดังกล่าวก็จะสอดคล้องกับวัตถุประสงค์การประเมิน และระยะเวลาหรือห้วงเวลาที่ทำการประเมิน โดยสรุปความสำคัญเป็น 5 ส่วน คือ

1. ความสำคัญของการประเมินต่อการวางแผนพัฒนาโครงการ ได้แก่ การประเมินก่อนดำเนินโครงการ การประเมินบริบทโครงการ การประเมินความต้องการจำเป็น (Needs assessment) ในส่วนนี้ ความสำคัญของการประเมิน ได้แก่ การทำให้มีการกำหนดวัตถุประสงค์ในการดำเนินงานที่ชัดเจน ทำให้แผนงานบรรลุตามวัตถุประสงค์ มีการวางแผนการใช้ทรัพยากรให้เกิดประโยชน์สูงสุด

2. ความสำคัญต่อการนำสารสนเทศมาเตรียมการ การประเมินความพร้อมของปัจจัยนำเข้า (Input evaluation) ปัจจัยใดที่เป็นปัญหาจะได้รับการจัดสรรให้อยู่ในจำนวนหรือปริมาณที่เหมาะสมเพียงพอแก่การดำเนินงาน ทรัพยากรที่ไม่จำเป็นหรือมีปริมาณมากเกินไปจะได้รับการตัดทอนและทรัพยากรใดขาดก็จะได้รับการจัดหาเพิ่มเติม

3. ความสำคัญต่อการนำสารสนเทศมาปรับปรุงโครงการ ได้แก่ การประเมินระหว่างดำเนินโครงการ (Formative evaluation) การประเมินกระบวนการ (Process evaluation) ความสำคัญในการประเมินกลุ่มนี้ ทำให้ได้สารสนเทศเกี่ยวกับปัญหาอุปสรรค ข้อดี ข้อเสียของการดำเนินโครงการ ทำให้มีแนวทางในการปรับปรุงแก้ไขโครงการ ลดความเสี่ยงของผลการดำเนินโครงการ ทำให้เกิดการควบคุมคุณภาพของงาน วิเคราะห์ทุกส่วนของโครงการ เพื่อไม่ให้เกิดปัญหาในการดำเนินโครงการ

4. ความสำคัญต่อการตัดสินใจผลการดำเนินโครงการ ได้แก่ การประเมินสรุปหรือการประเมินเมื่อสิ้นสุดโครงการ (Summative evaluation) การประเมินผลการดำเนินงาน (Product evaluation) การประเมินผลผลิต (Output evaluation) การประเมินโครงการในกลุ่มนี้มีความสำคัญในการทำให้ผู้บริหารได้สารสนเทศประกอบในการตัดสินใจ หากเป็นโครงการต่อเนื่องโครงการนี้ควรขยายหรือยุติโครงการ หรือไม่ อย่างไร หากไม่ใช่โครงการต่อเนื่องการประเมินในกลุ่มนี้จะมี ความสำคัญต่อการเป็นสารสนเทศในการดำเนินโครงการในลักษณะเดียวกันในโอกาสต่อไป

5. ความสำคัญต่อการติดตามผลของโครงการ ได้แก่ การประเมินติดตามผล (Follow up) การประเมินผลลัพธ์ (Outcome evaluation) การประเมินผลกระทบ (Impact evaluation) ทำให้ทราบข้อดี ข้อเสีย ของโครงการว่าผลผลิตของโครงการส่งผลต่อผู้เกี่ยวข้องอย่างไรในระยะยาว มีความสำคัญในการที่จะตัดสินใจระงับไปว่าโครงการใดมีประโยชน์มีคุณค่า หรือมีประโยชน์มากน้อยเพียงใด ตลอดจนแนวทางการวางแผนโครงการในวงรอบต่อไปที่ผู้บริหารจะใช้เป็นข้อมูลประกอบในการตัดสินใจว่าจะดำเนินการโครงการอย่างไร

สรุปความมุ่งหมายและความสำคัญการประเมินโครงการเพื่อช่วยเป็นข้อมูลในการประเมินคุณภาพของโครงการว่ามีประสิทธิภาพมากน้อยเพียงใด เป็นแนวทางในการปรับปรุงโครงการให้เหมาะสมยิ่งขึ้น

#### **ประเภทของการประเมินโครงการ**

นักวิชาการหลายท่านได้แบ่งประเภทการประเมินโครงการที่สำคัญ ดังนี้

ขวัญนภา อุณหกานต์ (2553, หน้า 13) จำแนกประเภทของการประเมินโครงการออกเป็น 4 ประเภทด้วยกัน คือ การประเมินก่อนดำเนินการ (Pre-evaluation) การประเมินขณะกำลังดำเนินการ (on-going evaluation) การประเมินเมื่อโครงการแล้วเสร็จ (End-of-project evaluation) และการประเมินหลังเสร็จสิ้นโครงการ (post-project evaluation)

พงศ์เทพ จิระโร (2554, หน้า 9-31) การประเมินโครงการสามารถจำแนกได้เป็น 3 ประเภท คือ จำแนกตามระยะเวลาที่ประเมิน จำแนกตามจุดมุ่งหมายของการประเมิน และจำแนกตามเนื้อหา โดยมีรายละเอียด ดังนี้

1. ประเมินจำแนกตามระยะเวลาที่ประเมิน เป็นการนำช่วงเวลามากำหนดประเภทของการประเมินโครงการ ชุดของช่วงเวลาที่นำมากำหนดประเภทที่พบได้โดยทั่วไป ได้แก่

1.1 การประเมินก่อนเริ่มโครงการ (Preparing evaluation) เป็นการประเมินเพื่อให้ได้สารสนเทศที่เกี่ยวข้องกับการดำเนินงาน โครงการ เพื่อศึกษาถึงความพร้อมของบุคคลและทรัพยากรที่เกี่ยวข้อง ประสิทธิภาพของปัจจัยนำเข้า ความเหมาะสมของกระบวนการที่คาดว่าจะนำมาใช้ในการบริหารจัดการโครงการ ความสมเหตุสมผลของการวางแผนดำเนินโครงการ รวมทั้งความเป็นไปได้ของโครงการ นอกจากนี้ยังมีจุดมุ่งหมายเพื่อตัดสินใจเลือกโครงการที่เหมาะสม จุดมุ่งหมายจะวิเคราะห์ข้อมูลก่อนดำเนินโครงการ เช่น การประเมินความต้องการ (Needs assessment) และศึกษาหาความเป็นไปได้ (Feasibility) คือ พิจารณาสภาพความพร้อมในการดำเนินโครงการ สารสนเทศที่ได้จะนำมาใช้ในการวางแผนการดำเนินงานของโครงการ รวมถึงวางแผนการประเมินขั้นต่อ ๆ ไป

1.2 การประเมินในระหว่างดำเนินโครงการ (Formative evaluation) เป็นการประเมินการดำเนินงานในด้านต่าง ๆ เมื่อนำโครงการที่วางแผนไว้ไปปฏิบัติ เพื่อตรวจสอบการดำเนินงานตามแผน ความก้าวหน้า ปัญหาอุปสรรค นำไปแก้ไขปรับปรุงให้การดำเนินงานเป็นไปอย่างมีประสิทธิภาพ เป็นการประเมินเพื่อทำให้ทราบปัญหาอุปสรรคแต่ละช่วง และนำผลการประเมินมาปรับปรุงโครงการ

1.3 การประเมินเมื่อเสร็จสิ้นโครงการหรือการประเมินผลสรุป (Summative evaluation) เป็นการประเมินเพื่อตอบคำถามว่าโครงการประสบผลสำเร็จตามวัตถุประสงค์หรือไม่ เป็นการประเมินแบบมุ่งผลสัมฤทธิ์ ตัวบ่งชี้ที่ใช้ประเมินจะมีความสอดคล้องกับวัตถุประสงค์โดยให้ความสนใจในผลการดำเนินโครงการ (Product) ที่เป็นผลผลิต (Output) เช่น ความรู้ของผู้เข้าร่วมอบรม จำนวนผู้สำเร็จ ระดับความพึงพอใจของผู้เข้าอบรม เป็นต้น เพื่อใช้เป็นข้อมูลสำหรับผู้รับผิดชอบโครงการว่าโครงการบรรลุตามแผนหรือไม่ สมควรจะทำโครงการนั้นต่อไปหรือไม่อย่างไร

2. ประเมินจำแนกตามจุดมุ่งหมายของการประเมิน โดยทั่วไปจำแนกออกเป็น 2 ประเภท ได้แก่ การประเมินเพื่อปรับปรุงพัฒนาหรือเพื่อพิจารณาความก้าวหน้าของโครงการ และการประเมินโครงการเพื่อตัดสินใจ แต่ในความต้องการผลการประเมินที่มีอยู่จริงในการดำเนินโครงการนอกจาก 2 จุดมุ่งหมายดังกล่าวแล้ว อาจพบการประเมินที่มีจุดมุ่งหมายเพื่อการวางแผนการดำเนินโครงการ การประเมินโครงการเพื่อจัดระดับผลงาน ในที่นี้จะเสนอจุดมุ่งหมายการประเมิน 2 กลุ่ม ดังนี้

2.1 การประเมินเพื่อปรับปรุงหรือเพื่อพัฒนาโครงการ สอดคล้องกับโมเดล การประเมินที่มีการประเมินความก้าวหน้า (Formative evaluation) การประเมินกระบวนการ (Process evaluation) มีจุดมุ่งหมายในการประเมินเพื่อนำผลมาเป็นสารสนเทศในการปรับปรุงโครงการ การประเมินกลุ่มนี้ทำให้ทราบว่าหลังจากดำเนินโครงการ ได้ระยะหนึ่ง โครงการได้มีความก้าวหน้าเพียงใด เป็นไปตามแผนหรือกำหนดการหรือไม่ มีกิจกรรมใดหรือมีส่วนใดของโครงการบ้างที่ต้องมีการปรับปรุง ทั้งนี้เนื่องจากกระบวนการดำเนินงานอาจมีหลายกิจกรรม บางกิจกรรมประสบความสำเร็จ แต่บางกิจกรรมอาจมีความเสี่ยงที่ต้องรับการปรับปรุง

2.2 การประเมินที่มีจุดมุ่งหมายเพื่อการตัดสินใจ เป็นการตัดสินใจสิ่งที่มุ่งประเมินนั้นว่า ได้หรือตก ตัดสินว่าผ่านหรือไม่ผ่าน ตัดสินว่ารับรองหรือไม่รับรอง บรรลุหรือไม่บรรลุ โดยเกณฑ์ตัดสินจะมีเพียง 2 ระดับ ข้อสังเกต คือ บางตัวบ่งชี้อาจกำหนดเกณฑ์ตัดสินได้เลย เช่น ตัวบ่งชี้ของผู้เข้าอบรม วัดเป็นคะแนน คิดเป็นร้อยละ เกณฑ์ตัดสินร้อยละ 50 ดังนั้น หากมีผลการวัดความรู้ต่ำกว่าร้อยละ 50 ตัดสินว่าตก หากมีผลการวัดมากกว่าร้อยละ 50 ตัดสินว่าได้

3. ประเมินจำแนกตามเนื้อหาสาระของการประเมิน เนื้อหาสาระที่แตกต่างกันสามารถนำมาใช้แบ่งประเภทของการประเมินโครงการได้ อาจแบ่งตามเนื้อหาสาระโดยการกำหนดประเภทตามแบบจำลอง (Model) การประเมินแบบต่าง ๆ ขึ้นอยู่กับการกำหนดวัตถุประสงค์ในการประเมินว่ามีวัตถุประสงค์การประเมินว่าอย่างไร อาจใช้ตามกรอบการประเมิน (Evaluation framework) ที่นำโมเดลการประเมินมาเป็นกรอบในการกำหนดหรืออาจเป็นการกำหนดขึ้นโดยผู้ประเมิน หรือกำหนดโดยเจ้าของโครงการที่ต้องการสารสนเทศหรือองค์ความรู้จากการประเมิน

พิชิต ฤทธิ์จรูญ (2555, หน้า 11) การประเมินโครงการแบ่งออกได้เป็นหลายประเภทแล้วแต่เกณฑ์ที่ใช้แบ่ง ในที่นี้ใช้เกณฑ์ในการแบ่งประเภท 4 เกณฑ์ คือ แบ่งตามลำดับเวลา การบริหารโครงการ แบ่งตามวัตถุประสงค์ของโครงการ แบ่งตามสิ่งที่ถูกประเมิน และแบ่งตามหลักที่ยึดในการประเมิน โดยมีรายละเอียด ดังนี้

1. แบ่งตามลำดับเวลาการบริหารโครงการ ได้แบ่งการประเมินเป็น 3 ประเภท คือ

1.1 การประเมินก่อนเริ่มโครงการหรือก่อนนำโครงการไปปฏิบัติ (Intrinsic evaluation) เป็นการประเมินที่อยู่ในขั้นตอนของการวางแผนโครงการ โดยมีจุดมุ่งหมายเพื่อการวิเคราะห์ข้อมูลต่าง ๆ ก่อนลงมือดำเนินการจัดทำโครงการ ซึ่งอาจทำได้ใน 2 ลักษณะ คือ

1.1.1 การศึกษาความเหมาะสมของโครงการ โดยการพิจารณาปัญหาและความจำเป็นในการจัดทำโครงการ อาจใช้เทคนิคของการสำรวจความต้องการ (Needs assessment) หรืออาจจะใช้เทคนิคของการศึกษาความเป็นไปได้ (Feasibility study) ซึ่งเป็นการพิจารณาสภาพความพร้อมของการจัดทำโครงการในด้านต่าง ๆ ที่จะทำให้มีความเป็นไปได้ในการปฏิบัติจริง

1.1.2 การวิเคราะห์โครงการหรือการประเมินร่างโครงการ (Project appraisal or analysis) เป็นการพิจารณาเอกสารโครงการเพื่อตรวจสอบความเหมาะสมของสิ่งที่กำหนด ความสอดคล้องระหว่างหัวข้อต่าง ๆ ตั้งแต่ หลักการและเหตุผล วัตถุประสงค์ เป้าหมาย วิธีดำเนินการ ฯลฯ เพื่อให้โครงการที่เสนอมีความสมบูรณ์มากยิ่งขึ้น

1.2 การประเมินขณะดำเนินงาน (Ongoing evaluation) เป็นการประเมินที่จัดทำ ในระหว่างการดำเนินโครงการ โดยมีจุดมุ่งหมายเพื่อการศึกษาความก้าวหน้าของโครงการ เป็นช่วง ๆ และเพื่อดูว่าการดำเนินโครงการได้เป็นไปตามแผนที่กำหนดไว้หรือไม่ มีปัญหาหรือ อุปสรรคอย่างไรในแต่ละช่วงของการทำงาน ข้อมูลสารสนเทศที่จะได้นำมาใช้เพื่อการปรับปรุง การดำเนินโครงการเพื่อให้ความเหมาะสมมากยิ่งขึ้นในช่วงระยะต่อไป การประเมินขณะ ดำเนินงานนี้ต้องอาศัยกลไกของระบบการกำกับติดตามงานมาช่วย จึงจะทำให้การประเมินเป็นไป อย่างมีประสิทธิภาพ

1.3 การประเมินหลังการดำเนินงาน (Pay-off evaluation) เป็นการประเมินที่จัดทำขึ้น เมื่อการดำเนินโครงการเสร็จสิ้นลงแล้ว โดยมีจุดมุ่งหมายเพื่อตรวจสอบดูว่า เมื่อได้ดำเนิน โครงการจนสิ้นสุดแล้ว ได้รับความสำเร็จตามวัตถุประสงค์และเป้าหมายที่กำหนดไว้มากน้อย เพียงใด ตลอดจนการติดตามตรวจสอบผลกระทบหรือผลข้างเคียงว่ามีอะไรเกิดขึ้นตามมาอีกบ้าง ถ้ากล่าวโดยสรุป การประเมินหลังการดำเนินงาน มีจุดเน้น ดังนี้

1.3.1 เป็นการประเมินผลที่เกิดขึ้น (Outcome) และผลกระทบ (Impact) ของ โครงการ

1.3.2 เป็นการประเมินผลรวมสรุปของโครงการทั้งหมด (Summative evaluation) คือ การประเมินตั้งแต่การวางแผนปฏิบัติงาน ค่าใช้จ่าย และผลผลิตของโครงการ

2. แบ่งตามวัตถุประสงค์ของการประเมิน ได้แบ่งการประเมินเป็น 2 ประเภท คือ

2.1 การประเมินความก้าวหน้าของโครงการ (Formative evaluation) เป็นการประเมิน ขณะทีโครงการกำลังดำเนินการอยู่ โดยมีมุ่งตรวจสอบ ควบคุม กำกับดูแลการดำเนินงาน ตลอดจน การศึกษาความก้าวหน้า ปัญหา อุปสรรค และข้อบกพร่องต่าง ๆ ในระหว่างการดำเนินโครงการ เพื่อนำผลการประเมินมาใช้ในการปรับปรุงแก้ไขการดำเนินงานให้เป็นไปด้วยความราบรื่น บรรลุตามวัตถุประสงค์และเป้าหมายของโครงการ

2.2 การประเมินผลรวมสรุป (Summative evaluation) เป็นการประเมินหลังจาก เสร็จสิ้นโครงการแล้ว เป็นการประเมินผลสัมฤทธิ์ของโครงการเพื่อมุ่งตรวจสอบว่า โครงการ ประสบผลสำเร็จตามวัตถุประสงค์หรือเป้าหมายตามที่กำหนดไว้หรือไม่ เพียงใด ผลการประเมิน

จะเป็นประโยชน์ต่อการตัดสินใจว่าควรยุติโครงการ ดำเนินโครงการต่อไป หรือปรับขยายโครงการ ในช่วงระยะต่อไป

### 3. แบ่งตามสิ่งที่ถูกประเมิน ได้แบ่งการประเมินเป็น 4 ประเภท คือ

3.1 การประเมินบริบทหรือสภาวะแวดล้อม (Context evaluation) เป็นการประเมินเกี่ยวกับนโยบาย เป้าหมาย สภาพเศรษฐกิจและสังคม ปัญหาและความต้องการของบุคคลและหน่วยงานที่เกี่ยวข้องกับโครงการว่า มีความสอดคล้องเอื้อต่อการจัดทำโครงการหรือไม่ ตลอดจนทรัพยากรและข้อจำกัดต่าง ๆ ในการดำเนินโครงการ สารสนเทศที่ได้นำมาใช้ประกอบการตัดสินใจเพื่อกำหนดวัตถุประสงค์และเป้าหมายของโครงการให้มีความสอดคล้องกับนโยบายของหน่วยงานระดับสูงกว่าและบริบทหรือสภาพขององค์กรที่จะจัดทำโครงการ

3.2 การประเมินปัจจัยเบื้องต้น (Input evaluation) เป็นการประเมินความพร้อมทั้งในเชิงคุณภาพและความพอเพียงของทรัพยากรต่าง ๆ ก่อนเริ่มโครงการว่า มีทรัพยากรพร้อมที่จะดำเนินโครงการได้หรือไม่ สารสนเทศที่ได้นำมาใช้ในการตัดสินใจเกี่ยวกับวิธีการของการใช้ทรัพยากรต่าง ๆ เพื่อให้การดำเนินโครงการสามารถบรรลุตามวัตถุประสงค์ที่กำหนดไว้

3.3 การประเมินกระบวนการ (Process evaluation) เป็นการประเมินขณะดำเนินงานหรือประเมินความก้าวหน้าของโครงการ สารสนเทศที่ได้นำมาใช้ในการตัดสินใจเพื่อการปรับปรุงการดำเนินโครงการ

3.4 การประเมินผลผลิตหรือผลงาน (Product evaluation) เป็นการประเมินหลังจากการดำเนินโครงการสิ้นสุดแล้ว ประกอบด้วยการประเมินผลลัพธ์ (Output evaluation) โดยพิจารณาจากปริมาณและคุณภาพของผลผลิตเปรียบเทียบกับวัตถุประสงค์ของโครงการ และการประเมินผลกระทบ (Impact evaluation) โดยเปรียบเทียบกับวัตถุประสงค์ของแผนงานหรือองค์การ สารสนเทศที่ได้นำมาใช้ในการตัดสินใจคุณค่าของผลผลิตของโครงการ ทั้งในด้านปริมาณและคุณภาพเพื่อการตัดสินใจว่า ควรจะคงโครงการไว้ ปรับขยายหรือยกเลิกโครงการ

### 4. แบ่งตามหลักยึดในการประเมิน ได้การประเมินเป็น 2 ประเภท คือ

4.1 การประเมินที่อิงวัตถุประสงค์/ เป้าหมายของโครงการหรืองานเป็นเกณฑ์ (Goal based evaluation) เป็นการประเมินที่ตัดสินคุณค่าของโครงการจากการเปรียบเทียบผลของโครงการกับผลที่คาดหวังตามวัตถุประสงค์หรือเป้าหมายของโครงการ

4.2 การประเมินที่ไม่อิงวัตถุประสงค์/ เป้าหมายของโครงการ (Goal-free evaluation) เป็นการประเมินที่ไม่จำเป็นต้องกำหนดเป้าหมายหรือวัตถุประสงค์ของการประเมินให้สอดคล้องกับเป้าหมาย หรือวัตถุประสงค์ของโครงการ กล่าว คือ การตัดสินคุณค่าของโครงการควรเน้นที่

การตีค่าของผลทั้งหมดที่เกิดขึ้น (Actual outcomes) จากโครงการ ซึ่งประกอบด้วยการประเมินผลที่คาดหวังตามเป้าหมายหรือวัตถุประสงค์ของโครงการและผลที่มีได้คาดหวัง

สรุปประเภทของการประเมินสามารถแบ่งออกได้เป็น 3 ระยะ ได้แก่ ประเมินก่อนดำเนินโครงการซึ่งเป็นการประเมินที่มีจุดมุ่งหมายเพื่อการตัดสินใจเกี่ยวกับการกำหนดและเลือกทำโครงการ ประเมินระหว่างดำเนินการเป็นการประเมินที่เกิดขึ้นในช่วงระยะเวลาของการปฏิบัติงาน เพื่อตรวจสอบการดำเนินโครงการว่าเป็นไปตามวัตถุประสงค์หรือไม่ มีปัญหาและอุปสรรคอะไรบ้าง และประเมินเมื่อสิ้นสุดโครงการเป็นการประเมินที่จัดทำขึ้นเพื่อประเมินโครงการว่าบรรลุจุดมุ่งหมายที่กำหนดไว้เพียงใด

### รูปแบบของการประเมินโครงการ

ในการวิจัยเชิงประเมินโดยทั่วไปมีหลายรูปแบบ เช่น รูปแบบของสเตค (Stake) สตัฟเฟิลบีม (Stufflebam) ไทเลอร์ (Tyler) โพรวัส (Provovus) เป็นต้น ซึ่งแต่ละรูปแบบจะแตกต่างกันออกไปตามความเหมาะสมของโครงการที่จะประเมิน ซึ่งนักประเมินจะเป็นผู้เลือกพิจารณาแบบตามความต้องการ เพื่อให้ได้ผลการประเมินตามจุดประสงค์ที่ตั้งไว้ นอกจากนี้ยังขึ้นอยู่กับความรู้ความสามารถและความเข้าใจเกี่ยวกับการใช้รูปแบบต่าง ๆ ของผู้ประเมินด้วย ได้มีผู้กล่าวถึงรูปแบบการประเมินโครงการไว้หลายท่าน ดังนี้

#### รูปแบบการประเมินของสตัฟเฟิลบีม (Stufflebeam)

Stufflebeam (1967) ได้เสนอแนวคิดเกี่ยวกับรูปแบบการประเมิน เรียกว่า ชิปโมเดล (CIPP Model) เป็นการประเมินที่เป็นกระบวนการต่อเนื่อง โดยมีจุดเน้นที่สำคัญ คือ ใช้ควบคู่กับการบริหารโครงการ เพื่อหาข้อมูลประกอบการตัดสินใจอย่างต่อเนื่องตลอดเวลา วัตถุประสงค์การประเมิน คือ การให้สารสนเทศเพื่อการตัดสินใจคำว่า CIPP เป็นคำย่อมาจากคำว่า Context input process และ Product Stufflebeam ได้ให้ความหมายว่า การประเมินเป็นกระบวนการของการบรรยาย การเก็บข้อมูล การวิเคราะห์ข้อมูลข่าวสาร เพื่อนำข้อมูลไปใช้ประโยชน์ในการตัดสินใจ เลือกทางเลือกที่เหมาะสม ซึ่งการประเมินเพื่อให้ได้สารสนเทศที่สำคัญ มุ่งประเมิน 4 ด้าน คือ การประเมินสภาพแวดล้อม (Context evaluation) การประเมินปัจจัยเบื้องต้น (Input evaluation) การประเมินกระบวนการ (Process evaluation) และการประเมินผลผลิต (Product evaluation) ประเภทการประเมินผลและลักษณะของการตัดสินใจ

#### รูปแบบการประเมินของ เคริกแพตทริก (Kirkpatrick)

Kirkpatrick (1975) แห่งมหาวิทยาลัยวิสคอนซินสหรัฐอเมริกา อดีตประธาน ASTD (The American Society for Training and Development) ได้นำเสนอแนวคิดเกี่ยวกับการฝึกอบรมและการประเมินผลการฝึกอบรมว่า “การฝึกอบรมเป็นการช่วยเหลือบุคลากรให้สามารถปฏิบัติงาน



ได้อย่างมีประสิทธิภาพ ในการฝึกอบรมใด ๆ ควรจะจัดให้มีการประเมินผลการฝึกอบรมซึ่งถือเป็นสิ่งจำเป็นที่จะช่วยให้รู้ว่าการจัดโปรแกรมการฝึกอบรมมีประสิทธิภาพเพียงใด”

Kirkpatrick เห็นว่าการประเมินผลการอบรมจะทำให้ได้ความรู้อย่างน้อย 3 ประการ คือ

1. การฝึกอบรมนั้นได้ทำอะไร หรือเกิดประโยชน์ต่อหน่วยงานในลักษณะใดบ้าง
2. ควรยุติโครงการชั่วคราวก่อน หรือควรดำเนินการต่อไปเรื่อย ๆ
3. ควรปรับปรุงหรือพัฒนาโปรแกรมการฝึกอบรมในส่วนใดบ้างอย่างไร

ขวัญภา อุณหานต์ (2553, หน้า 15-16) รูปแบบการประเมินเกิดขึ้นจากการที่นักประเมินต่างพยายามมุ่งที่เสนอสูตรสำเร็จต่าง ๆ ในการดำเนินการประเมินโครงการ ทั้งนี้เพื่อความสะดวกและสามารถเลือกใช้ได้อย่างเหมาะสม

กำจัด สุกโต (2553, หน้า 16) ได้สรุปรูปแบบการประเมินโครงการว่า มีหลากหลายรูปแบบแต่ละรูปแบบแตกต่างกันออกไปตามความเหมาะสมกับโครงการที่จะประเมิน จึงเป็นหน้าที่ของผู้ประเมินในการเลือกรูปแบบในการประเมินให้เหมาะสมกับโครงการ เพื่อให้ได้ผลการประเมินที่มีคุณค่าเชื่อถือได้ต่อไป

จากที่ได้กล่าวมาแล้วนั้นสามารถสรุปได้ว่า รูปแบบการประเมิน (Evaluation model) เป็นกรอบความคิดหรือแบบแผนที่แสดงให้เห็นถึงขั้นตอนหรือกระบวนการของการประเมิน ซึ่งอาจมีความแตกต่างกันทั้งนี้ขึ้นอยู่กับข้อตกลงเบื้องต้นของการประเมิน ซึ่งจะมีประโยชน์อย่างยิ่งต่อการวางแผนการประเมิน โดยทั่วไปนิยมแบ่งออกเป็น 3 รูปแบบ คือ

1. รูปแบบการประเมินที่เน้นการตัดสินใจ (Decision-oriented evaluation model) เป็นรูปแบบการประเมินที่มีจุดมุ่งหมายเพื่อให้ได้มาซึ่งข้อมูล สารสนเทศสำหรับใช้ในการตัดสินใจเกี่ยวกับโครงการนั้น ๆ ได้แก่ รูปแบบการประเมินของ Alkin, Stufflebeam, Hammond and Welch
2. รูปแบบการประเมินที่เน้นการตัดสินใจคุณค่า (Judgmental evaluation Models) เป็นรูปแบบการประเมินที่มีจุดมุ่งหมายเพื่อให้ได้มาซึ่งข้อมูล สารสนเทศสำหรับกำหนดและวินิจฉัยคุณค่าของโครงการนั้น ๆ ได้แก่ รูปแบบการประเมินของ Scriven, Stake and Provus
3. รูปแบบการประเมินที่เน้นจุดมุ่งหมาย (Objective-centered evaluation models) เป็นรูปแบบการประเมินที่เน้นการตรวจสอบผลที่เกิดขึ้นจากโครงการว่าบรรลุตามจุดมุ่งหมายที่กำหนดไว้มากน้อยเพียงใด ได้แก่ รูปแบบการประเมินของ Tyler, Cronbach and Kirkpatrick

### ประโยชน์ของการประเมินโครงการ

สมคิด พรหมจ้อย (2550, หน้า 37) ได้กล่าวถึงประโยชน์ของการประเมินโครงการ ซึ่งพอจะสรุปได้ ดังนี้

1. ช่วยให้อัปเดต และสารสนเทศต่าง ๆ เพื่อนำไปใช้ในการตัดสินใจเกี่ยวกับการวางแผนโครงการตรวจสอบความพร้อมของทรัพยากรต่าง ๆ ที่จำเป็นในการดำเนินโครงการ ตลอดจนตรวจสอบความเป็นไปได้ในการจัดกิจกรรมต่าง ๆ
2. ช่วยทำให้การกำหนดวัตถุประสงค์ของโครงการมีความชัดเจน
3. ช่วยในการจัดหาข้อมูลเกี่ยวกับความก้าวหน้าปัญหา และอุปสรรคของการดำเนินโครงการ
4. ช่วยให้อัปเดตเกี่ยวกับความสำเร็จ และความล้มเหลวของโครงการ เพื่อนำไปใช้ในการตัดสินใจ และวินิจฉัยว่าจะดำเนินโครงการในช่วงต่อไปหรือไม่ จะยกเลิกหรือขยายการดำเนินงานโครงการต่อไป
5. ช่วยให้ได้ข้อมูลที่บ่งบอกถึงประสิทธิภาพของการดำเนินโครงการว่าเป็นอย่างไร คำนวณค่ากับการลงทุนหรือไม่
6. เป็นแรงจูงใจให้ผู้ปฏิบัติงานในโครงการ เพราะการประเมินโครงการด้วยตนเอง จะทำให้ผู้ปฏิบัติงานได้ทราบผลการดำเนินงาน จุดเด่น จุดด้อย และนำข้อมูลไปใช้ในการปรับปรุงและพัฒนาโครงการให้มีประสิทธิภาพยิ่งขึ้น

สรุปประโยชน์ของการประเมินสามารถช่วยในการตัดสินใจในการนำโครงการไปใช้ การดำเนินโครงการควรมีต่อไปหรือต้องปรับปรุงโครงการ

### การประเมินโครงการตามรูปแบบของเคิร์กแพตทริก

Kirkpatrick (1975) ได้เสนอแนวคิดเกี่ยวกับการฝึกอบรมและการประเมินการฝึกอบรมว่า “การฝึกอบรมเป็นการช่วยเหลือบุคลากรให้สามารถปฏิบัติงานได้อย่างมีประสิทธิภาพ ในการฝึกอบรมใด ๆ ควรจัดให้มีการประเมินผลการฝึกอบรมซึ่งถือเป็นสิ่งจำเป็นที่จะช่วยให้รู้ว่าการจัดโปรแกรมการฝึกอบรมมีประสิทธิภาพเพียงใด” การฝึกอบรมเป็นกิจกรรมปกติที่เกิดขึ้นในทุกโครงการ เป็นกิจกรรมที่จัดขึ้นเพื่อพัฒนาบุคลากรในหน่วยงาน โดยมุ่งหวังให้ผู้ผ่านการอบรมได้มีการปรับปรุงเปลี่ยนแปลงแนวทางการทำงานให้มีประสิทธิภาพมากขึ้น และเห็นว่าการประเมินผลการอบรมจะทำให้ได้ความรู้อย่างน้อย 3 ประการ คือ

ประการที่หนึ่ง การฝึกอบรมนั้นได้ให้อะไรหรือเกิดประโยชน์อะไรต่อหน่วยงาน ในลักษณะใดบ้าง

ประการที่สอง ควรยุติโครงการชั่วคราว หรือควรดำเนินการต่อไปเรื่อย ๆ  
 ประการที่สาม ควรปรับปรุงหรือพัฒนา โปรแกรมการฝึกอบรมในส่วนใดบ้าง อย่างไร  
 Kirkpatrick ได้พัฒนารูปแบบการประเมินประสิทธิผลโครงการอบรม (Training model)  
 ที่เน้นการประเมินหลังโครงการอบรม 4 ระดับ ดังนี้

1. การประเมินปฏิกิริยาตอบสนอง (Reaction evaluation)
2. การประเมินการเรียนรู้ (Learning evaluation)
3. การประเมินพฤติกรรม (Behavior evaluation)
4. การประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation)

#### **การประเมินปฏิกิริยาตอบสนอง (Reaction evaluation)**

เพื่อให้รู้ว่าผู้เข้าร่วมการฝึกอบรมนั้นมีความรู้สึกอย่างไรต่อการฝึกอบรม วิธีการที่จะช่วยให้ได้รับข้อมูลเกี่ยวกับปฏิกิริยาตอบสนองที่มีความหมายและตรงตามความเป็นจริงจากผู้เข้ารับการอบรม ได้แก่

1. กำหนดให้แน่อนชัดเจนลงไปว่าต้องการได้รับข้อมูลอะไร เช่น ปฏิกิริยาตอบสนองของเนื้อหาในหลักสูตรการฝึกอบรม วิธีการฝึกอบรม วิทยากร สถานที่ฝึกอบรม ระยะเวลา บรรยากาศการฝึกอบรม เป็นต้น
2. ออกแบบเครื่องมือ หรือแบบสอบถามที่จะใช้ในการเก็บข้อมูล
3. ข้อคำถามที่ใช้ควรเป็นชนิดที่เมื่อได้รับข้อมูลหรือได้คำตอบแล้วสามารถนำคำตอบมาแปลงเป็นตัวเลขแจกแจงความถี่ หรือวิเคราะห์เชิงปริมาณได้
4. กระตุ้นให้ผู้รับการฝึกอบรมได้เขียนแสดงความคิดเห็นหรือข้อเสนอแนะเพิ่มเติมในข้อคำถามต่าง ๆ
5. เพื่อให้ผู้เข้ารับการฝึกอบรมได้แสดงปฏิกิริยาตอบสนองผ่านแบบสอบถามตามความเป็นจริง จึงไม่ควรให้ผู้เข้ารับการฝึกอบรมเขียนชื่อตนเองในแบบสอบถามและควรใช้เวลาอย่างเพียงพอที่จะให้คำตอบครบทุกข้อ

#### **แนวทางในการประเมินปฏิกิริยา**

1. ความคิดเห็นเกี่ยวกับความรู้ที่ได้รับจากโครงการหรือหลักสูตรการฝึกอบรมว่าผู้เข้ารับการฝึกอบรม ได้รับความรู้ ความเข้าใจในระดับใด (มาก ปานกลาง น้อย) มีความรู้สึกต่อหลักสูตรอย่างไรมีทักษะในการปฏิบัติเพิ่มขึ้นมากน้อยเพียงใด
2. ความคาดหวังที่จะได้รับและการนำความรู้และทักษะที่ได้รับจากฝึกอบรมไปใช้ในการทำงาน

3. ความคิดเห็นต่อเนื้อหา/ วิทยากรที่บรรยายมีความเหมาะสมกับเนื้อหาในการบรรยาย  
 มากน้อยเพียงใด

4. ความคิดเห็นต่อสถานที่ โสตทัศนูปกรณ์ ระยะเวลาในการฝึกอบรมว่า  
 มีความเหมาะสมมากน้อย เพียงใด

### การประเมินการเรียนรู้ (Learning evaluation)

มีวัตถุประสงค์เพื่อจะให้ผู้เข้ารับการฝึกอบรมได้รับความรู้และทักษะอะไรบ้าง  
 มีเจตคติอะไรบ้างที่เปลี่ยนแปลงไปจากเดิม ทั้งนี้เพราะความรู้ ทักษะ และเจตคติ ส่วนเป็น  
 องค์ประกอบพื้นฐานสำคัญที่จะช่วยให้เกิดการเปลี่ยนแปลงพฤติกรรมการทำงานของผู้เข้ารับ  
 การฝึกอบรมในอนาคตต่อไป โดยที่ต้องวัดทั้งความรู้ ทักษะ และเจตคติของผู้เข้ารับการฝึกอบรม  
 ทั้งก่อนและหลังฝึกอบรม และวิเคราะห์ทั้งคะแนนรายข้อและคะแนนรวม โดยการเปรียบเทียบ  
 ระหว่างก่อนและหลังการฝึกอบรม

#### แนวทางการประเมินการเรียนรู้

สามารถแยกเทคนิคและวิธีในการประเมิน ให้เป็นไปตามวัตถุประสงค์ที่ต้องการได้ ดังนี้

1. ความรู้/ ความคิด (Cognitive) ซึ่งเป็นการวัดความสามารถทางสมอง เช่น ความจำ  
 ความเข้าใจ การนำไปใช้ การวิเคราะห์ การสังเคราะห์ เทคนิคและวิธีการที่ใช้ก็จะเป็นแบบทดสอบ  
 ข้อเขียน ทั้งก่อนและหลังการฝึกอบรม หรือจากแบบฝึกหัด

2. เจตคติ (Attitude) ซึ่งเป็นการวัดด้านความรู้สึก เช่น ความสนใจความรู้สึก ค่านิยม  
 สามารถใช้วิธีการวัดได้หลายวิธี คือ ใช้แบบสอบถามเพื่อประเมินตนเอง บทบาทสมมติ พฤติกรรม  
 ในระหว่างเข้ารับการฝึกอบรม

3. ทักษะ (Skill) ซึ่งเป็นการวัดความสามารถในการปฏิบัติ เช่น ทักษะการพูด การเขียน  
 การปฏิบัติหรือการเคลื่อนไหวต่าง ๆ สามารถใช้การทดสอบการปฏิบัติงานของแต่ละบุคคล เช่น  
 หลังจากการฝึกอบรมการผลิตสื่อการสอนแล้ว ให้ผู้เข้ารับการฝึกอบรมมาปฏิบัติให้ดู  
 การใช้เครื่องมือวัดความรู้ ทักษะและเจตคติ มีอยู่ 2 แบบด้วยกัน คือ

3.1 ใช้แบบสอบวัดความรู้ ทักษะและเจตคติ ที่เป็นแบบวัดมาตรฐาน

3.2 สร้างแบบทดสอบขึ้นใช้เอง อาจใช้รูปแบบอย่างใดอย่างหนึ่ง หรือหลายอย่าง  
 หรือทุกอย่างต่อไปนี้ได้ เช่น

3.2.1 แบบ “ถูก” หรือ “ผิด”

3.2.2 แบบเห็นด้วย หรือ ไม่เห็นด้วย อาจเป็นมาตราส่วนประมาณค่า

3.2.3 แบบเลือกตอบที่เห็นว่าถูกต้องที่สุด

3.2.4 แบบเติมคำ/ ข้อความลงในช่องว่าง

### การประเมินพฤติกรรม (Behavior evaluation)

การประเมินพฤติกรรมมีวัตถุประสงค์เพื่อตรวจสอบว่า ผู้เข้ารับการฝึกอบรม ได้มีการเปลี่ยนแปลงพฤติกรรมในการทำงานไปในทิศทางที่พึงประสงค์หรือไม่ การประเมินพฤติกรรมมีวัตถุประสงค์เพื่อรู้ว่า เมื่อได้รับการฝึกอบรมไปแล้ว ผู้เข้ารับการฝึกอบรม ได้มีการเปลี่ยนแปลงพฤติกรรมในการทำงานไปในทิศทางที่พึงประสงค์หรือไม่ การประเมินผลในขั้นนี้นับว่ายากและใช้เวลามากกว่าการประเมินในสองขั้นแรก เพราะต้องออกไปติดตามประเมินผลในสถานที่ทำงานจริง ๆ ของผู้เข้ารับการฝึกอบรม ซึ่งมีคำถามอยู่หลายข้อที่ผู้ประเมินต้องตอบให้ได้ เช่น ควรจะออกไปประเมินผลเมื่อไร (1 เดือน 3 เดือน 6 เดือน หรือ 1 ปี ภายหลังจากการฝึกอบรม) จะเก็บข้อมูลจากใครจึงจะเชื่อถือได้มากที่สุด (จากผู้บังคับบัญชา จากเพื่อนร่วมงาน จากผู้ใต้บังคับบัญชา หรือจากผู้เข้าร่วมการฝึกอบรม) กำหนดว่ามีพฤติกรรมการทำงานอะไรบ้างที่คาดหวังจะให้เกิดการเปลี่ยนแปลงและเตรียมคำถามเพื่อเก็บข้อมูล โดยที่

1. ควรมีการวัดพฤติกรรมการทำงานของผู้เข้ารับการฝึกอบรม ทั้งก่อนและหลังการฝึกอบรม
2. ระยะเวลาระหว่างการฝึกอบรมกับการประเมินผลหลังการฝึกอบรมนั้น ควรจะห่างกันพอสมควร เพื่อให้แน่ใจว่าการเปลี่ยนแปลงพฤติกรรมการทำงานได้เกิดขึ้นจริง ๆ และควรมีการประเมินหลาย ๆ ครั้ง เป็นระยะ ๆ ไป เช่น ประเมินทุก ๆ 3 เดือน
3. ควรเก็บข้อมูลหลาย ๆ แหล่ง เช่น ผู้บังคับบัญชา เพื่อนร่วมงาน และจากกลุ่มผู้ที่ผ่านการอบรม

#### แนวทางการประเมินพฤติกรรม

1. ใช้แนวความคิดการประเมินแบบก่อนและหลังการฝึกอบรม คือ วัดพฤติกรรมก่อนที่ผู้เข้ารับการฝึกอบรมจะได้รับการถ่ายทอดความรู้จากวิทยากร แล้วเก็บผลของการวัดเพื่อเปรียบเทียบกับการวัดผลตอนสุดท้ายภายหลังเสร็จสิ้นการฝึกอบรม ก็จะทราบความแตกต่างระหว่างพฤติกรรมของผู้เข้ารับการฝึกอบรมว่า เปลี่ยนแปลงไปอย่างไร ทั้งนี้ควรจะใช้วิธีการสังเกตการสัมภาษณ์ ฯลฯ
2. การจัดตั้งมาตรฐานในการปฏิบัติงาน หรือเปรียบเทียบมาตรฐานคือหลังจากการฝึกอบรมไปแล้ว มีทักษะในการทำงานสูงขึ้นหรือไม่ เช่น ในการพิมพ์ดีดอาจตั้งมาตรฐานในการพิมพ์ดีดของพนักงานธุรการไว้ว่า จะต้องพิมพ์สัมผัสให้ได้นาทีละ 45 คำ เมื่อพนักงานธุรการมารับการฝึกอบรมแล้วนำผลการปฏิบัติงานไป เปรียบเทียบกับมาตรฐานสำหรับผลงานนี้

3. ผลการปฏิบัติงานหรือการกระทำ หรือพฤติกรรมการทำงานของผู้เข้ารับการฝึกอบรม จากการสังเกตของผู้บังคับบัญชา หรือหัวหน้างาน ผู้ใต้บังคับบัญชา และกลุ่มบุคคลอื่นที่คุ้นเคยกับการปฏิบัติงานของผู้เข้ารับการฝึกอบรม

4. การประเมินผลการปฏิบัติจากการประเมินงานประจำปี (ในรอบปีที่ผ่านมา)

5. ใช้การเปรียบเทียบการปฏิบัติงานระหว่างกลุ่มที่มีได้เข้ารับการฝึกอบรมกับกลุ่มที่ผ่านการฝึกอบรมว่าพฤติกรรมการทำงานทั้งสองกลุ่มแตกต่างกันหรือไม่เพียงใด การประเมินพฤติกรรม ในการประเมินตามแนวความคิดที่เสนอข้างต้น จะต้องใช้เวลา และอาศัยความชำนาญของผู้ประเมินเป็นอย่างมาก เพราะฉะนั้นควรประเมินกับโครงการฝึกอบรมขนาดใหญ่ หรือโครงการฝึกอบรมที่จัดต่อเนื่อง ส่วนโครงการขนาดเล็ก ควรใช้วิธีการง่าย ๆ คือ กำหนดว่ามีพฤติกรรมการทำงานอะไรบ้าง ที่คาดหวังจะให้เกิดการเปลี่ยนแปลง เตรียมคำถามที่ใช้ในการสัมภาษณ์ ทำการสัมภาษณ์บุคคลหลาย ๆ กลุ่ม ภายหลังจากสิ้นสุดการฝึกอบรมไประยะหนึ่ง เพื่อจะได้ทราบว่าพฤติกรรมที่คาดหวังเอาไว้เหล่านั้น ได้เกิดการเปลี่ยนแปลงจริงหรือไม่ ข้อมูลที่ได้จากการสัมภาษณ์ ควรเอามาแปลงตัวเลข แล้วทำการวิเคราะห์ในเชิงปริมาณ

#### **การประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation)**

การประเมินผลลัพธ์ที่เกิดต่อองค์กร เพื่อให้รู้ว่าการฝึกอบรมได้ก่อให้เกิดผลดีต่อหน่วยงานอย่างไรบ้าง เป็นการประเมินผลที่ต้องการจะให้เกิดขึ้นกับหน่วยงานซึ่งเป็นการประเมินผลที่ยากที่สุด เพราะในความเป็นจริงนั้นมีตัวแปรอื่นอีกมากมาย นอกเหนือจากการฝึกอบรมที่มีผลกระทบต่อหน่วยงานและตัวแปรเหล่านั้นบางทีก็ยากต่อการควบคุม ฉะนั้นอะไรก็ตามที่เกิดขึ้นแก่หน่วยงานในทางที่ดีจึงสรุปได้ยากว่าเป็นผลมาจากโปรแกรมการฝึกอบรม Kirkpatrick ได้ให้ข้อเสนอแนะในการประเมิน ผลลัพธ์ไว้ ดังนี้

1. ควรวัดสภาวะการณ์หรือเงื่อนไขต่าง ๆ ก่อนการฝึกอบรมเอาไว้ แล้วนำไปเปรียบเทียบกับสภาวะการณ์ภายหลังการฝึกอบรม โดยใช้ข้อมูลที่สังเกตได้หรือวัดได้

2. พยายามหาทางควบคุมตัวแปรอื่น ๆ ซึ่งคาดว่าจะมีอิทธิพลต่อการเปลี่ยนแปลงในผลที่ต้องการให้เกิดขึ้นกับหน่วยงาน วิธีหนึ่งที่จะทำได้ คือ การใช้กลุ่มควบคุมหรือกลุ่มทดลอง

#### **แนวทางในการประเมินผลลัพธ์**

1. ดัชนีบ่งชี้ของความพอใจของทีมงาน เช่น ลดอัตราการป่วย การลา ฯลฯ

2. ผลกระทบของโครงการฝึกอบรมที่ส่งผลกระทบต่อองค์กร เช่น ความจงรักภักดี ความภูมิใจที่มีต่อหน่วยงาน การทุ่มเทร่างกายและแรงใจในการทำงาน โดยมีได้หวังผลตอบแทนในรูปที่เป็นตัวเงินและไม่เป็นตัวเงิน ภาพพจน์ที่ดีของหน่วยงาน ต่อสายตาสมาชิกชน

3. ผลลัพธ์ของโครงการฝึกอบรมต่อองค์กรในด้านกำไร การเพิ่มประสิทธิภาพ  
ทางการบริหาร ฯลฯ

4. การวิเคราะห์ต้นทุน-ผลกำไร (จุดคุ้มทุน) ของโครงการฝึกอบรม

ข้อเสนอแนะในการประเมินผลลัพธ์ที่เกิดต่อองค์กร

1. ควรมีการวัดสภาพการณ์ หรือเงื่อนไขต่าง ๆ ก่อนการฝึกอบรม แล้วนำไป  
เปรียบเทียบกับสภาพการณ์ หรือเงื่อนไขภายหลังการฝึกอบรม โดยใช้ข้อมูลที่สังเกตได้หรือวัดได้

2. แนวทางในการควบคุมตัวแปรอื่น ๆ ที่คาดว่าจะมีอิทธิพลต่อการเปลี่ยนแปลง  
ต่อผลที่ต้องการให้เกิดขึ้นกับ องค์กร วิธีการหนึ่งที่ทำได้ก็คือ ใช้กลุ่มควบคุมกับกลุ่มทดลอง  
จากการศึกษารูปแบบในการประเมินโครงการฝึกอบรมแบบเคิร์กแพตทริก ซึ่งเป็นรูปแบบที่ใช้  
เป็นแนวทางในการประเมินโครงการ สรุปได้ว่า รูปแบบการประเมินโครงการฝึกอบรมแบบ  
เคิร์กแพตทริก เป็นรูปแบบที่เน้นการประเมินหลังโครงการอบรม มีการประเมินใน 4 ระดับ ดังนี้

2.1 การประเมินปฏิกิริยาตอบสนอง (Reaction evaluation) เพื่อให้รู้ว่าผู้เข้าอบรม  
พึงพอใจในโครงการหรือไม่

2.2 การประเมินการเรียนรู้ (Learning evaluation) เพื่อให้รู้ว่าผู้เข้าอบรม ได้เรียนรู้  
อะไรจากโครงการฝึกอบรมบ้าง

2.3 การประเมินพฤติกรรม (Behavior evaluation) เพื่อให้รู้ว่าผู้เข้าอบรมเปลี่ยน  
พฤติกรรมอันเนื่องมาจากสิ่งได้เรียนรู้มาหรือไม่

2.4 การประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) เพื่อให้รู้ว่าการเปลี่ยน  
พฤติกรรมมีผลกระทบในทางบวกต่อองค์กรหรือไม่

### การประเมินอภิมาน

การประเมินอภิมานเป็นการประเมินผลการประเมินที่มีความสำคัญโดยผู้วิจัยนำเสนอ  
การศึกษาเอกสารที่เกี่ยวข้องกับการประเมินอภิมาน 6 ด้าน คือ ความหมายของการประเมินอภิมาน  
จุดมุ่งหมายของการประเมินอภิมาน ประเภทของการประเมินอภิมาน หลักการประเมินอภิมาน  
ขั้นตอนของการประเมินอภิมาน และเกณฑ์การประเมินอภิมาน

#### ความหมายของการประเมินอภิมาน

ได้มีนักวิชาการทางการศึกษาให้ความหมายของการประเมินไว้ดังต่อไปนี้

สุวิมล ติรกานันท์ (2550, หน้า 10) ได้ให้ความหมายไว้ว่า การประเมินอภิมาน (Meta  
evaluation) หมายถึง การประเมินโครงการประเมินเพื่อตรวจสอบความถูกต้องของการประเมิน

นิยมใช้ประเมินการประเมินผลผลิต (Product evaluation) การประเมินผลสรุป (Summative evaluation) และการประเมินผลที่เกิดขึ้น (Outcome evaluation)

รัตน์ บัวสนธิ์ (2550, หน้า 10) ได้ให้ความหมายไว้ว่า การประเมินอภิมาน (Meta evaluation) หมายถึง กระบวนการดำเนินงานเพื่อให้ได้มาซึ่งสารสนเทศเชิงบรรยายและเชิงตัดสินใจเกี่ยวกับการมีประโยชน์ (Utility) ความเป็นไปได้ (Feasibility) ความเหมาะสม (Propriety) และความแม่นยำถูกต้อง (Accuracy) ของการประเมินตลอดจนสารสนเทศที่เกี่ยวกับการดำเนินงานประเมินอย่างเป็นระบบ ความเชื่อศรัทธา การยอมรับและความรับผิดชอบในการดำเนินงานประเมินที่มีต่อสังคม ทั้งนี้เพื่อนำผลที่ได้ไปเป็นแนวทางการดำเนินงานประเมิน รวมทั้งชี้แนะจุดอ่อนและจุดแข็งของรายการประเมิน

พิชญ์ ฟองศรี (2551, หน้า 12) ได้ให้ความหมายไว้ว่า การประเมินอภิมาน (Meta evaluation) หมายถึง การประเมินของการประเมินอีกครั้งหนึ่งเพื่อพิจารณาคุณภาพ โดยศึกษาความถูกต้องและความเหมาะสมของการประเมิน

สมหวัง พิธิยานุวัฒน์ (2553, หน้า 84) ให้ความหมายของการประเมินอภิมาน หมายถึง การตัดสินคุณค่าภายนอกและคุณค่าภายในของงานประเมิน

ปิยวรรณ กระจ่างงาม (2556, หน้า 29) ได้ให้ความหมายไว้ว่า การประเมินอภิมาน (Meta evaluation) เป็นกระบวนการตรวจสอบคุณภาพของงานประเมิน โดยอาศัยเกณฑ์ในการประเมินที่มีความครอบคลุมกระบวนการประเมินเพื่อทำให้ทราบถึงคุณภาพของงานประเมินและเป็นแนวทางในการปรับปรุงการประเมิน

สิริศักดิ์ อาจวิชัย (2556, หน้า 14-15) ได้ให้ความหมายไว้ว่า การประเมินอภิมานเป็นกลไกในการควบคุมคุณภาพประเมินงานประเมินหลาย ๆ เรื่อง หรือการประเมินผู้ประเมินหลาย ๆ คนเพื่อรวบรวมสารสนเทศจากงานประเมินไปใช้ในการตัดสินคุณค่า ความเหมาะสม ความเป็นไปได้ โดยเทียบกับเกณฑ์ (Criteria) และมาตรฐาน (Standard) ที่กำหนดไว้ ผลการประเมินจะทำให้นักประเมินได้สารสนเทศหรือข้อมูลที่ใช้ในการตัดสินใจยกเลิกหรือพัฒนาปรับปรุงงานประเมิน

เนติ เฉลยวาเรศ (2556, หน้า 2) กล่าวไว้ว่า การประเมินอภิมานเป็นการประเมินงานประเมินหรือคุณภาพของงานประเมินรวมไปถึงงานประเมินการสังเคราะห์วิธีที่ใช้ในการประเมินเพื่อให้ได้สารสนเทศที่เป็นประโยชน์ในการปรับปรุง และรายงานผลการประเมินในด้านจุดแข็งจุดอ่อนต่อสาธารณะ ซึ่งมีวิธีในการประเมินอภิมานแตกต่างกัน

จากความหมายของการประเมินอภิมาน สรุปได้ว่า เป็นกระบวนการประเมินเพื่อตรวจสอบพิจารณาคุณภาพของงานประเมิน โดยอาศัยเกณฑ์ในการประเมินที่มีความครอบคลุมกระบวนการประเมิน เพื่อนำผลที่ได้ไปปรับปรุง และเป็นแนวทางในการจัดกิจกรรมต่อไป



### จุดมุ่งหมายของการประเมินอภิมาน

นงลักษณ์ วิรัชชัย (2551, หน้า 5) ได้สรุปวัตถุประสงค์ของการประเมินอภิมานไว้ 2 ประการ คือ 1) เพื่อตรวจสอบคุณภาพ การควบคุม/ กำกับ และการรับรองคุณภาพของโครงการประเมินทั้งโครงการเดียวหรือหลายโครงการ หรือเป็นการประเมินศักยภาพของนักประเมินเพียงคนเดียวหรือหลายคน และ 2) เพื่อสังเคราะห์สรุปผลการประเมินกรณีที่มีการประเมินกรณีที่มีการประเมินหลายโครงการหรือมีผลการประเมินจากนักประเมินหลายคน ให้ได้ผลการประเมินนำไปใช้ประโยชน์ต่อไปในการพัฒนาทั้งคุณภาพการประเมิน นักประเมินและคุณภาพของหน่วยงาน

สิริศักดิ์ อาจิวชัย (2556, หน้า 15-16) ได้สรุปวัตถุประสงค์ในการประเมิน คือ เพื่อควบคุม ตรวจสอบ และรับรองคุณภาพงานประเมินหรือนักประเมิน และสังเคราะห์ผลการประเมินจากหลาย ๆ โครงการประเมิน โดยทำการประเมินการบรรลุเป้าหมาย ผลกระทบ ประสิทธิภาพและประสิทธิผลเพื่อนำเสนอข้อมูลสารสนเทศในการนำไปปรับปรุงพัฒนาต่อไป จากที่กล่าวมาจุดมุ่งหมายของการประเมินอภิมานเพื่อตรวจสอบควบคุมและรับรองคุณภาพของการประเมินโครงการ และเป็นกลไกอย่างหนึ่งในการควบคุมคุณภาพของการประเมินอันจะนำไปสู่การพัฒนาโครงการต่อไป

### ประเภทของการประเมินอภิมาน

สมหวัง พิธิยานุวัฒน์ (2553, หน้า 84-85) ได้กล่าวถึงประเภทของการประเมินอภิมานว่าสามารถแบ่งได้ออกเป็น 2 ประเภท กล่าว คือ การประเมินอภิมานเชิงความก้าวหน้า (Formative meta-evaluation) และการประเมินอภิมานแบบสรุปรวม (Summative meta-evaluation) โดยมีรายละเอียดดังต่อไปนี้

#### 1. การประเมินอภิมานเชิงความก้าวหน้า (Formative meta-evaluation)

การประเมินอภิมานเชิงความก้าวหน้าเป็นการประเมินงานที่มีลักษณะการใช้การประเมินเชิงสร้างสรรค์ กล่าว คือ เป็นการวางแผนก่อนปฏิบัติการประเมิน การประเมินอภิมานเชิงความก้าวหน้าช่วยให้นักประเมินตัดสินใจเกี่ยวกับการตัดสินใจว่าจะประเมินอะไร นิยามปัญหาที่ประเมิน กำหนดผู้เกี่ยวข้องในการประเมินหรือลูกค้าของการประเมิน ทำวัตถุประสงค์ของการประเมินให้กระจ่าง กำหนดข้อสัญญา กำหนดข้อมูลและสารสนเทศที่ต้องการ กำหนดวิธีการเก็บรวบรวมข้อมูล และวิธีวิเคราะห์ข้อมูล กำหนดตารางการประเมิน จัดหาทีมประเมิน กำหนดงบประมาณ เตรียมข้อสรุปและข้อเสนอแนะ และเลือกรูปแบบและวิธีการรายงานผลการประเมิน โดยการประเมินอภิมานในเชิงความก้าวหน้า แสดงดังภาพที่ 3

แผนการประเมินและ การดำเนินการประเมิน	การประเมินตามแนวทาง มาตรฐานที่กำหนด	การหลีกเลี่ยงข้อผิดพลาดหรือ หลุมพรางของการประเมิน
---	--	--

### ภาพที่ 3 สมการการประเมินอภิมานในเชิงความก้าวหน้า

โดยสรุปการประเมินอภิมานเชิงความก้าวหน้า หมายถึง เป็นการประเมินว่าแผนการประเมินและการดำเนินการประเมินเป็นไปตามแนวทางมาตรฐานที่กำหนดหรือไม่ ถ้าปฏิบัติตามแนวทางก็จะได้ผลการประเมินที่ถูกต้อง สมเหตุสมผล อีกทั้งเป็นการหลีกเลี่ยงข้อผิดพลาดโดยทั่วไปของการประเมินได้ ทั้งนี้ข้อตกลงเบื้องต้นที่สำคัญ คือ คุณค่าสูงสุดของการประเมินจะมีค่ามากที่สุดเมื่อดำเนินการประเมินตามแนวทางมาตรฐานของการประเมินที่กำหนดและหลีกเลี่ยงข้อผิดพลาดในการประเมิน

#### 2. การประเมินอภิมานเชิงสรุปรวม (Summative meta evaluation)

การประเมินอภิมานเชิงสรุปรวมถือว่าเป็นบทบาทพื้นฐานของการประเมินอภิมาน กล่าว คือ เป็นการแสดงถึงคุณค่ารวมของงานประเมินเป็นบทบาทที่แสดงถึงความพร้อมที่จะให้ตรวจสอบได้ของนักประเมิน โดยรายงานให้สาธารณชนทราบถึงความเป็นมาตรฐานของงานประเมิน ผลการประเมินอภิมานเชิงสรุปรวมช่วยให้ผู้ใช้ผลการประเมินตัดสินใจว่าจะยอมรับผลการประเมินและข้อเสนอแนะจากงานประเมินมากน้อยเพียงใด

#### หลักการประเมินอภิมาน

ศิริชัย กาญจนวาสิ (2550, หน้า 175-177) ได้เสนอไว้ว่า การประเมินอภิมานควรกำหนดเกณฑ์ตัดสินคุณภาพการประเมินให้ครอบคลุมการประเมินคุณภาพ 3 ด้าน คือ

1. คำถามการประเมิน ควรตรงประเด็นกับความต้องการจำเป็นที่แท้จริงของการจัดกิจกรรมหรือโครงการนั้น ๆ เพื่อใช้เป็นแนวทางในการตัดสินคุณค่าของสิ่งที่ต้องการให้เกิดขึ้นที่แท้จริงตามเป้าหมายของการจัดกิจกรรมหรือโครงการ ในขณะที่เดียวกันคำถามการประเมินควรสนองความต้องการการใช้สารสนเทศของผู้ที่เกี่ยวข้องสำหรับการตัดสินใจปรับปรุง/เปลี่ยนแปลง การจัดกิจกรรมหรือโครงการเพื่อพัฒนาสังคมต่อไปในอนาคต ดังนั้นการวิเคราะห์ถึงความต้องการจำเป็นที่แท้จริงและการมีส่วนร่วมในการกำหนดคำถามการประเมินของผู้ต้องการใช้สารสนเทศจึงมีความสำคัญต่อคุณภาพของคำถามการประเมิน

2. วิธีการประเมิน ควรเป็นกระบวนการที่ตอบสนองต่อจุดมุ่งหมายของการประเมิน สามารถจัดกระทำให้บรรลุผลของการตัดสินคุณค่าของสิ่งที่มีประเมินได้อย่างถูกต้องหรือ

เหมาะสม เนื่องจากคุณค่าของสิ่งที่มีมุ่งประเมินขึ้นอยู่กับบริบทการประเมินจึงควรมีลักษณะที่สอดคล้องและเหมาะสมกับบริบทตลอดจนเป็นที่เข้าใจและยอมรับได้ของผู้ที่เกี่ยวข้อง ดังนั้น จะเห็นได้ว่าองค์ประกอบสำคัญที่จะนำไปสู่การเลือกวิธีประเมินที่เหมาะสม ได้แก่ การกำหนดวัตถุประสงค์ และเป้าหมายของการประเมินที่ชัดเจน สภาพแวดล้อมของการประเมินประสบการณ์/ค่านิยมของนักประเมินและผู้ที่เกี่ยวข้อง

3. ผลการประเมิน ควรมีคุณลักษณะสำคัญที่สุด คือ สามารถนำไปใช้ประโยชน์ต่อผู้ที่เกี่ยวข้องอย่างกว้างขวาง ส่วนจะเป็นประโยชน์ต่อใครบ้างถือว่าเป็นหน้าที่ของนักประเมินที่ต้องสามารถระบุตั้งแต่เริ่มต้นทำการประเมินว่าใครบ้างที่อยู่ในข่ายจะเป็นผู้ใช้ผลการประเมิน เพื่อนักประเมินจะได้สร้างสัมพันธภาพที่ดีกับบุคคลเหล่านั้น จะได้ทราบว่าเขาต้องการข้อมูลทำไรบ้าง และเป็นแนวทางสู่ความร่วมมือด้านต่าง ๆ การให้ความสนใจ การยอมรับผลการประเมิน และการใช้ผลประเมินให้เป็นประโยชน์ ผลการประเมินควรมีความยุติธรรมในการตัดสินคุณค่าของสิ่งที่มีมุ่งประเมิน โดยจะต้องมีการเก็บรวบรวมข้อมูลอย่างรอบด้าน และจากผู้เกี่ยวข้องหลาย ๆ ฝ่ายมีการใช้เกณฑ์การตัดสินที่สมเหตุสมผลและยอมรับได้ พร้อมทั้งมีการเชื่อมโยงคุณค่าของสิ่งที่มีมุ่งประเมินภายใต้บริบทนั้นสู่คุณค่าของสิ่งนั้นในสายตาของสาธารณชน นอกจากความเป็นประโยชน์และความยุติธรรมแล้วผลการประเมินควรมีความน่าเชื่อถือ คือ สามารถเปิดเผยความจริงเกี่ยวกับคุณค่าของสิ่งที่มีมุ่งประเมินได้ ตัวบ่งชี้ที่สำคัญที่บ่งบอกถึงความน่าเชื่อถือของผลการประเมินอาจพิจารณาได้จากความน่าเชื่อถือของ “นักประเมิน” ในแง่ความเชี่ยวชาญในระเบียบวิธีการประเมิน เนื้อหา และจรรยาบรรณ ความสมเหตุสมผลของการกำหนดวัตถุประสงค์ของการประเมิน มีการออกแบบและใช้วิธีการประเมินที่เชื่อถือได้

#### ขั้นตอนของการประเมินอภิมาน

รัตนะ บัวสนธ์ (2555, หน้า 176-177) ได้สรุปจากแนวคิดของ Stufflebeam (2000, pp. 191-202) ถึงขั้นตอนกระบวนการการประเมินอภิมานว่าประกอบด้วย 11 ขั้นตอน กล่าวคือ

1. กำหนดและติดต่อกลุ่มผู้มีส่วนได้ส่วนเสียสำหรับการประเมินอภิมาน ซึ่งกลุ่มบุคคลดังกล่าว ได้แก่ ผู้ให้ผลการประเมิน ผู้ว่าจ้างทำการประเมินและผู้ปฏิบัติงานในโครงการที่ได้รับการประเมิน

2. จัดหาทีมนักประเมินอภิมานที่ต้องมีนักประเมินที่มีความเชี่ยวชาญอย่างน้อย 1 คน หรืออาจมากกว่านั้นก็ได้ร่วมกับบุคคลที่มีความเชี่ยวชาญในโครงการนั้น ๆ หรือบุคคลภายนอกองค์กร หน่วยงาน เจ้าของโครงการ ซึ่งเป็นที่ยอมรับหรือมีความน่าเชื่อถือทางสังคม บุคคลเหล่านี้เป็นส่วนประกอบที่สำคัญของการสร้างทีมนักประเมินอภิมานที่จะทำหน้าที่ประเมินการประเมินให้เป็นที่ยอมรับต่อผู้เกี่ยวข้องทั้งหลาย

3. กำหนดและนิยามคำถามการประเมินอภิमान โดยพิจารณาจากประเด็นคำถามหลัก 2 ประการ คือ 1) งานประเมินนั้นเป็นไปตามเงื่อนไขความต้องการหรือมาตรฐานการประเมินหรือไม่ หรือกล่าวง่าย ๆ ว่างานประเมินนั้นมีคุณค่าเพียงไรนั่นเอง และ 2) งานประเมินนั้นสอดคล้องกับความต้องการของผู้ใช้บริการหรือไม่ หรือเป็นงานประเมินที่มีคุณประโยชน์ต่อการนำไปใช้เพียงไร คำถามหลัก 2 คำถามนี้จะนำไปสู่การแตกเป็นประเด็นคำถามย่อย ๆ เพื่อใช้ในการประเมินอภิमान ซึ่งประเด็นคำถามย่อยสำหรับการประเมินอภิमानก็มักจะใช้หลักสำหรับนักประเมิน (Guiding principle for evaluation) และมาตรฐานการประเมินโครงการ (Program evaluation standards) ของคณะกรรมการกำหนดมาตรฐานการประเมิน รวมทั้งการกำหนดคำถามจากกลุ่มผู้ให้บริการด้วยส่วนหนึ่ง เหล่านี้ประกอบกันเป็นคำถามเพื่อใช้ในการดำเนินประเมินอภิमान

4. ยอมรับและใช้มาตรฐาน หลักการ และเกณฑ์ที่เหมาะสมสำหรับการตัดสินใจในการประเมิน ซึ่งความเหมาะสมของมาตรฐาน หลักการ และเกณฑ์ที่จะนำมาใช้นั้นก็อาจจะพิจารณาจากการยอมรับในวงวิชาการ วิชาชีพนั้น ๆ เช่น มาตรฐานความเป็นครู พยาบาล หรือวิศวกร เป็นต้น ขึ้นอยู่กับว่าเป็นการประเมินโครงการอะไรในแวดวงวิชาชีพใด

5. จัดทำสัญญายอมรับการว่าจ้างพร้อมเงื่อนไขข้อตกลงต่าง ๆ ในการประเมินอภิमान การทำบันทึกข้อตกลงสัญญาว่าจ้างการทำการประเมินอภิमानภายหลังจากการทำการเจรจาต่อรองระหว่างนักประเมินอภิमानและผู้บริการหรือผู้ว่าจ้างนั้นมีความสำคัญเป็นอย่างมาก ทั้งเป็นการช่วยป้องกันนักประเมินมิให้ถูกแทรกแซงในการเลือกเกี่ยวกับการประเมินแล้ว ยังช่วยให้ผู้รับบริการรวมถึงผู้ที่มีส่วนเกี่ยวข้องได้เข้าถึงการประเมินอภิमानซึ่งอาจจะต้องมีการเก็บรวบรวมข้อมูลต่าง ๆ เพิ่มเติม คู่มือประการหนึ่งเป็นการคอยตรวจสอบจับผิดนักประเมินโครงการเดิม ดังนั้น การทำสัญญาข้อตกลงที่ชัดเจนรัดกุมจะช่วยป้องกันความเข้าใจผิดในทำนองที่กล่าวได้เป็นอย่างดี

6. เก็บรวบรวมและตรวจสอบทบทวนสารสนเทศเกี่ยวกับการประเมินโครงการที่จัดเตรียมหรือปรากฏให้เห็นได้อย่างชัดเจน การเก็บรวบรวมและตรวจสอบสารสนเทศในขั้นนี้มุ่งหมายทำการตรวจสอบรายงานการประเมินโครงการที่ได้รับการจัดเตรียมนำเสนอหรือเผยแพร่ไว้แล้ว เช่น การเผยแพร่ในรูปเอกสาร สิ่งพิมพ์ หรืออาจเป็นการเผยแพร่ทางเว็บไซต์ (Website) ก็ได้ หากนักประเมิน พบว่า สารสนเทศที่ปรากฏอย่างชัดเจนซึ่งมีการจัดเตรียมไว้เหล่านี้ มีความสมบูรณ์ก่อให้เกิดความน่าเชื่อถือแล้วก็จะสามารถจัดทำเป็นรายงานการประเมินอภิमानได้เลย แต่ถ้าหากยังเห็นว่าสารสนเทศต่าง ๆ ที่มีอยู่ยังไม่เพียงพอต่อการสรุปจัดทำรายงานนักประเมินอภิमानก็ต้องดำเนินการต่อไปในขั้นที่ 7

7. เก็บรวบรวมข้อมูลใหม่ ๆ ที่จำเป็นเพิ่มเติมด้วยวิธีการต่าง ๆ เช่น การเยี่ยมชมสถานที่เพื่อสัมภาษณ์ การสังเกต การสำรวจ เป็นต้น ในการประเมินอภิमानเพื่อตัดสินคุณค่าและคุณประโยชน์ของการประเมินครั้งแรกนั้น บ่อยครั้งที่นักประเมินอภิमानจำเป็นต้องเก็บรวบรวมข้อมูลใหม่ ๆ เพิ่มเติมนอกเหนือไปจากข้อมูลและสารสนเทศที่จัดเตรียมไว้ในการประเมินครั้งแรก ซึ่งอาจยังไม่เพียงพอต่อการลงข้อสรุปตัดสินคุณค่าและคุณประโยชน์ของการประเมินนั้น ๆ การเก็บรวบรวมข้อมูลใหม่นี้อาจใช้วิธีการต่าง ๆ ได้แก่ การเยี่ยมชมสถานที่ที่โครงการดำเนินงานอยู่เพื่อการสัมภาษณ์ สังเกตบุคคลที่เกี่ยวข้องตลอดจนร่องรอยการดำเนินงานของโครงการหรืออาจจะใช้การสำรวจความคิดเห็นของบุคคลผู้มีส่วนได้ส่วนเสียก็ได้

8. ทำการวิเคราะห์ข้อมูลและสารสนเทศทั้งเชิงคุณภาพและเชิงปริมาณ เนื่องจากข้อมูลสารสนเทศทั้งที่มีอยู่เดิมจากการประเมินโครงการครั้งแรก และจากการเก็บรวบรวมข้อมูลใหม่ มีหลากหลายซึ่งอาจเป็นทั้งข้อมูลเชิงปริมาณและเชิงคุณภาพ ดังนั้นจึงต้องใช้วิธีการวิเคราะห์ทั้งสองแบบให้สอดคล้องกับลักษณะข้อมูลสารสนเทศที่มีอยู่

9. ตัดสินคุณค่าและคุณประโยชน์ของการประเมิน โดยใช้เกณฑ์มาตรฐานและหลักการตัดสินที่เหมาะสมเป็นที่ยอมรับได้ในวิชาการวิชาชีพนั้น ๆ

10. จัดทำรายงานการประเมินอภิमानตามที่จำเป็นหรือตามที่ระบุไว้ในสัญญาข้อตกลง รวมทั้งการนำเสนอรายงานการประเมินอภิमान การจัดทำรายงานการประเมินอภิमानเพื่อให้ตอบสนองกับความต้องการของผู้ใช้บริการ จำเป็นต้องมีการวางแผนเขียนรายงาน การจัดทำต้นฉบับทั้งที่เป็นรายงานย่อและรายงานสรุปสุดท้าย นอกจากนั้นเมื่อจัดทำรายงานแล้วก็อาจต้องมีการนำเสนอโดยวาจาประกอบในที่ประชุมของกลุ่มผู้ให้บริการก็ได้

11. ช่วยผู้รับบริการและผู้มีส่วนได้ส่วนเสียทำการตีความและนำผลการประเมินอภิमानไปใช้ในกรณีที่จำเป็นและหากเป็นไปได้นอกเหนือจากการทำรายงานและนำเสนอรายงานแล้ว นักประเมินอภิमानยังอาจจะช่วยตีความผลการประเมินอภิमानและเสนอแนะเกี่ยวกับการนำผลการประเมินไปใช้ในแง่มุมต่าง ๆ ให้แก่กลุ่มผู้มีส่วนได้ส่วนเสียหรือผู้รับบริการจากโครงการที่ทำการประเมินอีกด้วย

#### **เกณฑ์การประเมินอภิमान**

สิ่งสำคัญประการหนึ่งของการประเมินอภิमानก็คือ “เกณฑ์การประเมินอภิमान” ที่เป็นมาตรฐานที่กำหนดขึ้นเพื่อใช้ในการพิจารณาตัดสินคุณค่างานประเมินทั้งนี้เกณฑ์การประเมินอภิमानที่สำคัญมีอยู่ 2 แนวทาง ได้แก่ เกณฑ์การประเมินอภิमानตามแนวคิดของ Stufflebeam et al. และเกณฑ์การประเมินอภิमानตามแนวคิดของ Scriven โดยมีรายละเอียดของแต่ละแนวคิดดังต่อไปนี้

## 1. เกณฑ์การประเมินอภิमानตามแนวคิดของ Stufflebeam et al.

คณะกรรมการกำหนดมาตรฐานการประเมินทางการศึกษา (The Joint Committee on Standards for Educational evaluation) ซึ่งมี Stufflebeam เป็นผู้อำนวยการได้ดำเนินการจัดทำมาตรฐานสำหรับการตัดสินคุณภาพการประเมินทางการศึกษา ซึ่งแบ่งออกเป็น 4 มาตรฐาน คือ มาตรฐานด้านการใช้ประโยชน์ (Utility standards) มาตรฐานด้านความเป็นไปได้ (Feasibility standards) มาตรฐานด้านความเหมาะสม (Propriety standards) และมาตรฐานด้านความถูกต้อง (Accuracy standards) จำนวน 30 เกณฑ์โดยมีรายละเอียดดังต่อไปนี้ (Madaus, Scriven & Stufflebeam, 1983, pp. 395-404; Worthen & Sander, 1987, pp. 371-375; Popham, 1993, pp. 348-352; ศิริชัย กาญจนวาสี, 2545, หน้า 175-181 อ้างถึงใน ไชยยศ ไพโรพวิทย์ศิริธรรม, 2550, หน้า 63-68)

1.1 มาตรฐานด้านการใช้ประโยชน์ (Utility standards) เป็นมาตรฐานที่มีเจตนาเพื่อให้แน่ใจว่าการประเมินนั้นให้สารสนเทศที่จำเป็นแก่ผู้เกี่ยวข้อง ประกอบด้วยเกณฑ์จำนวน 8 ข้อ ดังนี้

U1 การระบุผู้เกี่ยวข้องที่ต้องการใช้สารสนเทศ: จะเป็นการระบุผู้เกี่ยวข้องหรือผู้ได้รับผลกระทบเกี่ยวกับการประเมินหรือผู้ต้องการใช้สารสนเทศในการประเมิน

U2 ความเป็นที่น่าเชื่อถือของผู้ประเมิน: ผู้ประเมินจะต้องเป็นผู้ที่รู้ถึงวิธีการประเมิน ผู้ประเมินจะต้องเป็นที่น่าเชื่อถือและได้รับการยอมรับตลอดจนมีความซื่อสัตย์ (น่าไว้วางใจ)

U3 การรวบรวมข้อมูลครอบคลุมและตอบสนองความต้องการใช้สารสนเทศของผู้เกี่ยวข้อง: การรวบรวมสารสนเทศจะต้องมีความครอบคลุมและเลือกใช้สารสนเทศที่ตรงกับวัตถุประสงค์การประเมิน ตลอดจนตอบสนองต่อความต้องการและความสนใจของผู้เกี่ยวข้อง

U4 การแปลความหมายและการตัดสินคุณค่ามีความชัดเจน: การแปลความหมายจากข้อค้นพบที่ได้จะต้องใช้มุมมองกระบวนการและเหตุผล โดยจะต้องบรรยายอย่างระมัดระวัง และการตัดสินคุณค่าจะต้องมีความยุติธรรม

U5 รายงานการประเมินมีความชัดเจนทุกขั้นตอน: รายงานการประเมินจะต้องบรรยายตั้งแต่วัตถุประสงค์ของการประเมิน บริบทของการประเมินความมุ่งหมาย กระบวนการ และข้อค้นพบของการประเมิน โดยจะต้องมีการอธิบายให้ผู้ที่เกี่ยวข้องกับผลการประเมินได้เข้าใจว่าประเมินอะไร ประเมินทำไม ใช้สารสนเทศอะไร บทสรุปเป็นอย่างไรและมีข้อเสนอแนะอะไรบ้าง

U6 การเผยแพร่ผลการประเมินไปยังผู้เกี่ยวข้องอย่างทั่วถึง: ข้อค้นพบที่ได้จากการประเมินจะต้องเผยแพร่ไปยังผู้เกี่ยวข้องกับการประเมิน โดยที่ผู้เกี่ยวข้องกับการประเมินสามารถประเมินและนำข้อค้นพบไปใช้ได้

U7 รายงานการประเมินเสร็จทันเวลาสำหรับการนำไปใช้ประโยชน์: รายงานการประเมินจะต้องเสร็จทันเวลาผู้เกี่ยวข้องกับการประเมินสามารถใช้สารสนเทศที่ได้จากรายงานการประเมินมากที่สุด

U8 การประเมินส่งผลกระทบต่อในการกระตุ้นให้มีการดำเนินการประเมินต่อไป  
อย่างต่อเนื่อง: การประเมินจะต้องกระตุ้นให้เกิดการวางแผนและปฏิบัติต่อไป

1.2 มาตรฐานด้านความเป็นไปได้ (Feasibility standards) เป็นมาตรฐานที่มีเจตนาเพื่อให้แน่ใจว่าการประเมินนั้นสอดคล้องกับความเป็นจริงสามารถปฏิบัติได้จริง มีความรอบคอบ ถ้วนถี่ประหยัดและคุ้มค่า ประกอบด้วยเกณฑ์จำนวน 3 ข้อ ดังนี้

F1 วิธีการประเมินสามารถนำไปปฏิบัติได้จริง: กระบวนการของการประเมินสามารถปฏิบัติได้จริง

F2 การเป็นที่ยอมรับได้ทางการเมือง: การออกแบบและการปฏิบัติการประเมินควรได้รับการยอมรับจากผู้เกี่ยวข้องทุกฝ่าย

F3 ผลที่ได้มีความคุ้มค่า: การประเมินควรได้รับสารสนเทศที่เพียงพอที่จะพิสูจน์ถึงความถูกต้อง เพื่อที่จะสามารถปรับปรุง ขยายงานหรือการล้มเลิก

1.3 มาตรฐานด้านความเหมาะสม (Propriety standards) เป็นมาตรฐานที่มีเจตนาเพื่อให้คำนึงถึงกฎ ระเบียบจรรยาบรรณ/ จรรยาบรรณ และความปลอดภัยของผู้ที่เกี่ยวข้องกับการประเมินประกอบด้วยเกณฑ์จำนวน 8 ข้อ ดังนี้

P1 การกำหนดข้อตกลงของการประเมินอย่างเป็นทางการ: มีการกำหนดข้อตกลงเบื้องต้นในการประเมิน (ทำไมถึงประเมิน ประเมิน อย่างไร ใครเป็นผู้ให้ข้อมูลประเมินเมื่อใด)

P2 การแก้ปัญหาของความขัดแย้งในการประเมินด้วยความเป็นธรรมและโปร่งใส: การให้ความสนใจเกี่ยวกับความขัดแย้ง (ที่เกิดขึ้นและหลีกเลี่ยงไม่ได้) และการปฏิบัติกับสิ่งที่เกิดขึ้นอย่างซื่อสัตย์และตรงไปตรงมา

P3 รายงานผลการประเมินอย่างตรงไปตรงมา เปิดเผยและคำนึงถึงข้อจำกัดของการประเมิน: การรายงานการประเมินจะต้องมีความตรงไปตรงมาซื่อสัตย์ เปิดเผย ตรงกับข้อค้นพบที่ได้จากการประเมิน และบอกถึงข้อจำกัดของการประเมินด้วย

P4 มีการให้ความสำคัญต่อสิทธิในการรับรู้ข่าวสารของสาธารณะ: การประเมินจะต้องให้ความสำคัญและรับรองสิทธิในการรับรู้ข่าวสาร โดยจะต้องคำนึงถึงความปลอดภัยและสิทธิส่วนบุคคล

P5 การคำนึงถึงสิทธิส่วนบุคคลของกลุ่มตัวอย่าง: การประเมินจะต้องออกแบบและปฏิบัติโดยคำนึงถึงสิทธิและความปลอดภัยของกลุ่มตัวอย่างทั้งนี้จะต้องให้ความสำคัญและปกป้องกลุ่มตัวอย่าง

P6 การเคารพสิทธิในการมีปฏิสัมพันธ์ของผู้เกี่ยวข้อง: ผู้ประเมินจะต้องเคารพสิทธิศักดิ์ศรีความเป็นมนุษย์ในการมีปฏิสัมพันธ์กับบุคคลอื่นที่เกี่ยวข้องกับการประเมิน

P7 มีรายงานผลการประเมินที่สมบูรณ์ ยุติธรรมและเสนอทั้งจุดเด่น จุดด้อยของสิ่งที่ประเมิน: รายงานการประเมินจะต้องมีความสมบูรณ์และยุติธรรมในการนำเสนอจุดเด่นและจุดด้อยของสิ่งที่ทำการประเมิน

P8 ผู้ประเมินทำการประเมินด้วยความรับผิดชอบและมีจรรยาบรรณ: ผู้ประเมินจะต้องใช้ทรัพยากรอย่างรอบคอบ มีความรับผิดชอบและมีจริยธรรม/ จรรยาบรรณ

1.4 มาตรฐานด้านความถูกต้อง (Accuracy standards) เป็นมาตรฐานที่มีเจตนาเพื่อให้แน่ใจว่าการประเมินเป็นไปอย่างเปิดเผย ตรงไปตรงมา มีการใช้เทคนิควิธีการเพื่อให้ได้สารสนเทศที่เพียงพอสำหรับการตัดสินใจคุณค่าทางด้านรูปธรรมและนามธรรมของสิ่งที่ทำการประเมิน ประกอบด้วยเกณฑ์จำนวน 11 ข้อ ดังนี้

A1 มีการระบุวัตถุประสงค์ประสงค์ของการประเมินอย่างชัดเจน: วัตถุประสงค์ของการประเมินจะต้องได้รับการตรวจสอบอย่างดี และจะต้องมีความชัดเจน

A2 มีการวิเคราะห์บริบทของการประเมินอย่างเพียงพอ: จะต้องมีการตรวจสอบรายละเอียดของบริบทการประเมินในรายงานการประเมิน

A3 มีการบรรยายจุดประสงค์และกระบวนการประเมินอย่างชัดเจน: วัตถุประสงค์และกระบวนการประเมินจะต้องมีการบรรยายอย่างละเอียด เพียงพอและสามารถปฏิบัติได้

A4 มีการบรรยายแหล่งข้อมูลและการได้มาอย่างชัดเจน: จะต้องมีการบรรยายแหล่งสารสนเทศ (ข้อมูล) ที่จะใช้อย่างละเอียด ชัดเจนและเพียงพอ

A5 การพัฒนาเครื่องมือและการเก็บรวบรวมข้อมูลที่มีความตรง: เครื่องมือและกระบวนการประเมินจะต้องมีความตรง และได้สารสนเทศที่ถูกต้องแม่นยำ

A6 การพัฒนาเครื่องมือและการเก็บรวบรวมข้อมูลที่มีความเที่ยง: เครื่องมือและกระบวนการประเมินจะต้องมีความเที่ยง และได้สารสนเทศที่มีความคงเส้นคงวา



A7 มีการจัดระบบควบคุมสำหรับการเก็บรวบรวมข้อมูล การวิเคราะห์และ  
การรายงาน: การรวบรวมข้อมูล การวิเคราะห์และการรายงานในการประเมินได้รับการพิจารณา/  
ตรวจสอบเพื่อให้ได้ผลการประเมินที่ไม่มีข้อบกพร่องผิดพลาด

A8 มีการวิเคราะห์ข้อมูลเชิงปริมาณ: มีการใช้การวิเคราะห์ข้อมูลเชิงปริมาณ  
ในการประเมิน เพื่อสนับสนุนการแปลความหมายของผลการประเมิน

A9 มีการวิเคราะห์ข้อมูลเชิงคุณภาพ: มีการใช้การวิเคราะห์ข้อมูลเชิงคุณภาพ  
ในการประเมิน เพื่อสนับสนุนการแปลความหมายของผลการประเมิน

A10 มีการลงสรุปที่มีเหตุผลและมีผลการประเมินสนับสนุน: การสรุปผลการประเมิน  
จะต้องมีความแน่นอน ชัดแจ้ง มีเหตุผลสนับสนุน

A11 มีการเขียนรายงานที่มีความเป็นปรนัย (Objectivity): การเขียนรายงาน  
การประเมินจะต้องมีความชัดเจนถูกต้องตามข้อค้นพบจากการประเมิน โดยปราศจากข้อคิดเห็น  
ที่ลำเอียงของผู้ประเมิน

นอกจากนี้ Stufflebeam, Scriven and Madaus (1983, p. 404) ได้ทำการวิเคราะห์  
ความสัมพันธ์ของเกณฑ์การประเมินอภิमान 4 มาตรฐาน 30 เกณฑ์ข้างต้นกับขั้นตอน/ ภาระงาน  
(Tasks) ในการประเมิน (ซึ่งประกอบด้วย 10 ขั้นตอน/ ภาระงาน ประกอบด้วย 1) การตัดสินใจ  
ที่จะทำการประเมิน (Deciding whether to do a study) 2) การกำหนดเป้าหมายการประเมิน (Clarity  
and assessing purpose) 3) การเป็นที่ยอมรับได้ทางการเมือง (Ensure political viability) 4) การทำ  
สัญญา (Contract) 5) การจัดทีมผู้ประเมิน (Staff the study) 6) การบริหารจัดการ/ ดำเนินการ  
ประเมิน (Manage the study) 7) การรวบรวมข้อมูล (Collect data) 8) การวิเคราะห์ข้อมูล (Analyze  
data) 9) การรายงานข้อค้นพบ (Report findings) และ 10) การนำผลการประเมินไปใช้ (Apply  
results) ) แสดงดังตารางที่ 1

ตารางที่ 1 ความสัมพันธ์ของเกณฑ์การประเมินอภิมานและขั้นตอนการประเมิน

เกณฑ์การประเมิน อภิมาน	ขั้นตอนการประเมิน									
	1. การตัดสินใจที่จะประเมิน	2. การกำหนดเป้าหมาย	3. การเป็นที่ยอมรับทางการเมือง	4. การทำสัญญา	5. การจัดทีมผู้ประเมิน	6. การจัดทำแผน การประเมิน	7. การรวบรวมข้อมูล	8. การวิเคราะห์ข้อมูล	9. การรายงานข้อค้นพบ	10. การนำผลการประเมินไปใช้
U1	✓	✓	✓	✓		✓			✓	✓
U2	✓		✓	✓	✓	✓	✓			
U3				✓			✓			
U4		✓	✓				✓	✓	✓	✓
U5						✓			✓	
U6			✓	✓		✓			✓	✓
U7				✓					✓	
U8	✓	✓	✓						✓	✓
F1			✓				✓	✓		
F2	✓		✓	✓	✓	✓	✓			✓
F3	✓	✓				✓				
P1	✓		✓	✓		✓	✓			✓
P2	✓	✓	✓	✓	✓	✓				✓
P3				✓					✓	
P4			✓	✓					✓	
P5			✓	✓		✓	✓			✓
P6			✓			✓	✓			
P7							✓		✓	✓
P8			✓	✓		✓				
A1	✓	✓		✓						
A2	✓	✓					✓	✓	✓	
A3	✓	✓	✓	✓		✓	✓	✓	✓	✓
A4							✓		✓	✓
A5							✓		✓	

ตารางที่ 1 (ต่อ)

เกณฑ์การประเมิน อภิमान	ขั้นตอนการประเมิน									
	1. การตัดสินใจที่จะประเมิน	2. การกำหนดเป้าหมาย	3. การเป็นที่ยอมรับทางการเมือง	4. การทำสัญญา	5. การจัดทีมผู้ประเมิน	6. การจัดดำเนิน การประเมิน	7. การรวบรวมข้อมูล	8. การวิเคราะห์ข้อมูล	9. การรายงานข้อค้นพบ	10. การนำผลการประเมินไปใช้
A6							✓			
A7						✓	✓			
A8								✓		
A9								✓		
A10		✓						✓	✓	✓
A11			✓		✓				✓	

จากตารางที่ 1 จะเห็นได้ว่าเกณฑ์การประเมินที่กำหนดโดยคณะกรรมการกำหนดมาตรฐานการประเมินทางการศึกษาทั้ง 4 มาตรฐาน 30 เกณฑ์นั้นมีความครอบคลุมภาระงานของการประเมินในทุกขั้นตอนเริ่มตั้งแต่การตัดสินใจที่จะประเมินการกำหนดเป้าหมายของการประเมิน การเป็นการเป็นที่ยอมรับได้ทางการเมือง การทำสัญญา การจัดทีมผู้ประเมิน การบริหารจัดการ/ ดำเนินการประเมิน การรวบรวมข้อมูล การวิเคราะห์ข้อมูล การรายงานข้อค้นพบและการนำผลการประเมินไปใช้

## 2. เกณฑ์การประเมินอภิमानตามแนวคิดของสคริฟเวน

Scriven (1991, pp. 230-231) ได้เสนอรายการตรวจสอบการประเมินหลัก (The key evaluation checklist: KEC) เพื่อใช้ในการประเมินคุณภาพงานประเมินในการประเมินอภิमान โดยแบ่งออกเป็น 13 ประเด็น ซึ่งครอบคลุมประเด็นการตรวจสอบตั้งแต่ก่อนเริ่มต้นการประเมินจนกระทั่งการวิเคราะห์ผลการประเมิน โดยมีรายละเอียดของรายการตรวจสอบการประเมิน ดังนี้

2.1 การบรรยาย (Description) เป็นการบรรยายเกี่ยวกับสิ่งที่จะประเมิน การออกแบบการประเมิน การกำหนดทีมประเมิน หน้าที่ที่จะต้องปฏิบัติ การกำหนดระยะเวลาในการประเมิน รวมถึง แหล่งข้อมูลที่จะใช้ในการเก็บรวบรวมข้อมูล

2.2 ภูมิหลังและบริบท (Background and context) เป็นการระบุผู้ต้องการทราบผลการประเมิน รวมถึงผู้ที่ไม่เห็นด้วยกับการประเมิน

2.3 ผู้เกี่ยวข้องกับการประเมิน (Consumer) เป็นการระบุถึงผู้เกี่ยวข้อง ผู้รับฟัง/รับทราบ/อ่านผลของการประเมิน

2.4 แหล่งที่มาของการประเมิน (Resources) เป็นการระบุถึงแหล่งที่จะได้มาซึ่งสารสนเทศในการประเมิน

2.5 คุณค่าของการประเมิน (Values) เป็นการระบุถึงคุณค่าสำหรับการประเมิน เช่น การตรงต่อเวลา (ทันเวลา) ความคุ้มค่า ความยุติธรรม ความตรง (Validity) ความเหมาะสม และตรงกับความต้องการของผู้ต้องการทราบผลการประเมิน

2.6 กระบวนการในการประเมิน (Process) เป็นการระบุถึงการออกแบบและวิธีการปฏิบัติในการประเมินที่จะต้องมีความเหมาะสมและถูกต้อง โดยอาจจะพิจารณาจากการที่มีการอำนวยความสะดวกในการประเมิน การรายงานและเผยแพร่ผลการประเมิน การพิจารณาถึงความเป็นส่วนตัวของผู้ให้ข้อมูล ความน่าเชื่อถือและความรับผิดชอบของผู้ประเมิน การพิจารณาถึงผลกระทบที่เกิดจากการประเมิน

2.7 ผลลัพธ์ของการประเมิน (Outcomes) เป็นการรายงานผลดีและผลเสียของการประเมินในขอบเขตซึ่งสามารถคาดการณ์ล่วงหน้า

2.8 ความคุ้มค่าของการประเมิน (Costs) เป็นการพิจารณาค่าใช้จ่ายในการประเมิน ทั้งที่เป็นตัวเงินและไม่ใช่ว่าเงิน

2.9 การเปรียบเทียบผลการประเมิน (Comparisons) เป็นการพิจารณาเปรียบเทียบผลการประเมิน ซึ่งอาจจะเป็นการตัดสินอย่างไม่เป็นทางการหรือการใช้โมเดลอื่น ๆ มาช่วยในการเปรียบเทียบ

2.10 การสรุปอ้างอิงผลการประเมิน (Generalizability) เป็นการพิจารณาถึงการที่ผู้อื่นสามารถนำผลของการประเมินไปใช้ในการออกแบบสำหรับการประเมินต่อไป

2.11 ความสำคัญของการประเมิน (Significance) เป็นการพิจารณาโดยการสังเคราะห์เกณฑ์การประเมินทุกข้อในข้างต้น โดยการพิจารณาผลที่ได้จากการอ่านที่เกี่ยวกับคุณค่าของงานประเมิน

2.12 ข้อเสนอแนะในการประเมิน (Recommendation) เป็นการพิจารณาถึงข้อเสนอแนะสำหรับผู้อ่านงานประเมิน ทั้งนี้อาจจะมีสิ่งที่แนะนำหรือสิ่งที่ไม่แนะนำให้ทำตาม เช่น ปฏิเสธให้ทำตามให้มีการทำซ้ำเพื่อยืนยันหรือนำไปดำเนินการต่อ เป็นต้น

### 2.13 การรายงาน (Report) เป็นการรายงานผลการประเมิน

เมื่อพิจารณาเกณฑ์การประเมินอภิमानทั้ง 2 แนวคิดข้างต้น พบว่า เกณฑ์การประเมินอภิमानตามแนวคิดของStufflebeam et al. และแนวคิดของ Scriven มีความสอดคล้องและคล้ายคลึงกันเป็นส่วนใหญ่ซึ่งเกณฑ์การประเมินอภิमानต่างก็มีความครอบคลุมกระบวนการประเมินตั้งแต่ก่อนเริ่มมีการประเมินจนกระทั่งถึงขั้นตอนของการนำผลการประเมินที่ได้ไปใช้ประโยชน์ แต่ที่ พบว่า เกณฑ์การประเมินอภิमानตามแนวคิดของ Stufflebeam et al. จะมีความละเอียดมากกว่า เกณฑ์การประเมินของ Stufflebeam อาจเนื่องมาจากเกณฑ์การประเมินตามแนวคิดของ Scriven มีการกล่าวถึงเกณฑ์การประเมินอภิमानไว้เป็นประเด็นอย่างกว้าง ๆ แต่ครอบคลุมสาระการประเมิน ในขณะที่เกณฑ์การประเมินอภิमानตามแนวคิดของ Stufflebeam et al. มีการแบ่งเกณฑ์การประเมินอภิमानออกเป็นรายมาตรฐาน และภายใต้มาตรฐานแต่ละด้าน ได้มีเกณฑ์การประเมินที่กล่าวถึงรายละเอียดของการประเมินไว้อย่างครอบคลุมขั้นตอนการประเมินเพื่อให้เห็นถึงความสอดคล้องสัมพันธ์กันของเกณฑ์การประเมินอภิमानทั้ง 2 แนวคิดผู้วิจัยได้ทำการวิเคราะห์ออกเป็นตารางเมตริกซ์แสดงความสัมพันธ์ของเกณฑ์การประเมินอภิमानทั้ง 2 แนวคิด ดังตารางที่ 2

ตารางที่ 2 ความสัมพันธ์ของเกณฑ์การประเมินอภิमानตามแนวคิดของ Stufflebeam et al. และแนวคิดของ Scriven

แนวคิดของ Stufflebeam	แนวคิดของ Scriven												
	การบรรยาย	ภูมิหลังและบริบท	ผู้เกี่ยวข้อง	แหล่งที่มา	คุณค่า	กระบวนการ	ผลลัพธ์	ความคุ้มค่า	การเปรียบเทียบ	การสรุปอ้างอิง	ความสำคัญ	ข้อเสนอแนะ	การรายงาน
U1		✓	✓										
U2	✓												
U3	✓												
U4					✓	✓			✓				
U5													✓
U6										✓			
U7	✓				✓					✓			
U8							✓				✓	✓	
F1						✓							

ตารางที่ 2 (ต่อ)

แนวคิดของ Stufflebeam	แนวคิดของ Scriven												
	การบรรยาย	ภูมิหลังและบริบท	ผู้เกี่ยวข้อง	แหล่งที่มา	คุณค่า	กระบวนการ	ผลลัพธ์	ความคุ้มค่า	การเปรียบเทียบ	การสรุปอ้างอิง	ความสำคัญ	ข้อเสนอแนะ	การรายงาน
F2		✓	✓			✓							
F3					✓			✓		✓			
P1					✓								
P2						✓							
P3					✓		✓						
P4						✓							
P5						✓							
P6						✓							
P7						✓	✓				✓	✓	
P8						✓							
A1	✓												
A2	✓	✓											
A3	✓												
A4	✓			✓									
A5						✓							
A6						✓							
A7						✓							
A8						✓							
A9						✓							
A10							✓				✓		
A11						✓							✓

ในครั้งนีผู้วิจัยเลือกใช้เกณฑ์การประเมินตามหลักการที่เป็นแนวทางสำหรับนักประเมิน และเกณฑ์การประเมินอภิมานการประเมินโครงการที่ Stufflebeam พัฒนาขึ้น โดยอิงกับมาตรฐานการประเมินโครงการในปี ค.ศ. 1999 (รัตนะ บัวสนธิ์, 2555) มาใช้ในการประเมินโครงการสร้าง

ข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการประเมินของเคิร์กแพททริกและการประเมินอภิमान

## แบบทดสอบ

### ความหมายของแบบทดสอบ

มาเรียม นิลพันธุ์ (2553, หน้า 164) แบบทดสอบหรือข้อทดสอบ หรือแบบสอบ หมายถึง ชุดของข้อคำถาม (Items) หรือชุดของสิ่งเร้าที่นำไปกระตุ้นให้บุคคลเกิดการตอบสนองในรูปแบบต่าง ๆ เช่น การเลือกตอบ การเขียนตอบ การแสดงพฤติกรรมให้พูด แสดงท่าทาง เพื่อให้สามารถวัดได้ สังเกตได้ ซึ่งจะนำไปสู่การแปลความหมาย

สรุปแบบทดสอบ คือ ชุดของคำถาม ปัญหา สถานการณ์ กลุ่มของงานหรือกิจกรรมอย่างใดอย่างหนึ่งที่ใช้เป็นสิ่งเร้า กระตุ้นช่วย หรือชักนำให้ผู้ถูกทดสอบแสดงพฤติกรรมหรือปฏิกิริยาตอบสนองตามแนวทางที่ต้องการ แบบทดสอบเป็นเครื่องมือวัดสมรรถภาพทางสมองได้ดีที่สุด

### ประเภทของแบบทดสอบ

สมนึก ภัททิยธนี (2551, หน้า 62) กล่าวว่า ประเภทของแบบทดสอบ สามารถแบ่งได้หลายลักษณะขึ้นอยู่กับเกณฑ์ที่ใช้ในการแบ่ง ดังนี้

#### 1. แบ่งตามสิ่งที่วัด แบ่งออกได้ 3 ประเภท คือ

1.1 แบบทดสอบวัดผลสัมฤทธิ์ (Achievement test) หมายถึง ข้อสอบที่วัดสมรรถภาพสมองด้านต่าง ๆ ที่นักเรียนได้รับการเรียนรู้ผ่านมาแล้วว่ามีอยู่เท่าใด แบ่งออกเป็น 2 ชนิด คือ

1.1.1 แบบทดสอบที่ครูสร้างขึ้น (Teacher made test) เป็นข้อสอบมุ่งวัดผลสัมฤทธิ์ของผู้เรียนเฉพาะกลุ่มที่ครูสอนจะไม่นำไปใช้กับนักเรียนกลุ่มอื่นเป็นแบบทดสอบที่ใช้กันทั่ว ๆ ไปในโรงเรียน

1.1.2 แบบทดสอบมาตรฐาน (Standardized test) เป็นข้อสอบที่มุ่งวัดผลสัมฤทธิ์เช่นเดียวกับแบบทดสอบที่ครูสร้างขึ้น แต่มีจุดมุ่งหมายเพื่อเปรียบเทียบคุณภาพต่าง ๆ ของนักเรียนที่ต่างกลุ่มกัน เช่น เปรียบเทียบคุณภาพของนักเรียนใน โรงเรียนแห่งหนึ่งกับนักเรียนกลุ่มอื่น ๆ ทั่วประเทศ (แบบทดสอบมาตรฐานระดับชาติ) หรือกับนักเรียนกลุ่มอื่น ๆ ทั่วจังหวัด (แบบทดสอบมาตรฐานระดับจังหวัด) เป็นต้น

ข้อคำถามของแบบสอบมาตรฐานจะมีลักษณะเช่นเดียวกับแบบทดสอบที่ครูสร้างขึ้น แต่ที่ต่างกัน คือ แบบทดสอบมาตรฐานต้องกำหนดมาตรฐานในการดำเนินการสอบ และการตรวจ

ให้คะแนนเป็นอย่างเดียวกัน และที่ต่างกันอย่างเห็นเด่นชัดก็คือ มีเกณฑ์ปกติ (Norms) สำหรับเป็นมาตรฐานในการเปรียบเทียบ เพื่อแปลความหมายของคะแนน

1.2 แบบทดสอบวัดความถนัด (Aptitude test) หมายถึง ข้อสอบที่มุ่งวัดสมรรถภาพสมองของผู้เรียนว่าจะสามารถเรียนต่อไปหรือประสบผลสำเร็จเพียงใด เพื่อใช้ในการพยากรณ์หรือทำนายอนาคตของผู้เรียน โดยอาศัยข้อเท็จจริงในปัจจุบันเป็นพื้นฐาน แบบทดสอบวัดความถนัดแบ่งออกเป็น 2 ชนิด คือ

1.2.1 แบบทดสอบวัดความถนัดทางการเรียน (Scholastic aptitude test) หมายถึง แบบทดสอบที่มุ่งวัดความสามารถทางวิชาการต่าง ๆ เช่น ภาษา คณิตศาสตร์ เป็นต้น แบบทดสอบนี้ใช้วัดเพื่อพยากรณ์ว่าผู้เรียนแต่ละคนจะสามารถเรียนต่อในทางใดได้สำเร็จ

1.2.2 แบบทดสอบวัดความถนัดเฉพาะ (Specific aptitude test) หมายถึง แบบทดสอบที่มุ่งวัดความถนัดเฉพาะที่เกี่ยวกับงานอาชีพต่าง ๆ หรือความสามารถพิเศษ เช่น ความสามารถทางดนตรี ศิลปะ เครื่องยนต์ การประดิษฐ์ เป็นต้น

1.3 แบบทดสอบวัดบุคลิกภาพและทางสังคม (Personal and social test) หมายถึง แบบทดสอบที่ใช้วัดบุคลิกภาพ (Personality) และการปรับตัว (Adjustment) ให้เข้ากับสังคม ซึ่งเป็นเรื่องที่ยาก ผลที่ได้ไม่คงที่แน่นอน เนื่องจากความเปลี่ยนแปลงในตัวบุคคลและสังคม และมักจะเรียกแบบสอบชนิดนี้ว่า แบบวัด (Scale) เพราะเป็นสมรรถภาพที่เกี่ยวข้องพัวพันระหว่างด้านพุทธิพิสัย (Cognitive domain) กับด้านจิตพิสัย (Affective domain) มีข้อสังเกตว่า แบบทดสอบ แบบวัด และแบบสอบถาม มีลักษณะแตกต่างกัน ดังนี้

1.3.1 แบบทดสอบ (Test) เมื่อตอบแต่ละข้ออาจจะถูก (ได้คะแนน) หรือผิด (ไม่ได้คะแนน)

1.3.2 แบบวัด (Scale) ใช้วัดเกี่ยวกับคุณลักษณะบางอย่าง มีความลึกซึ้งกว่าแบบสอบถามผลการตอบแต่ละข้อจะได้คะแนนแน่นอน (ไม่มีการตอบผิด) แต่คะแนนจะแตกต่างกันตามระดับที่กำหนดให้ เช่น คะแนน 5 ระดับ เป็น 1 2 3 4 หรือ 5 คะแนน เป็นต้น

1.3.3 แบบสอบถาม (Questionnaire) ใช้วัดเกี่ยวกับความคิดเห็น (Opinion) หรือความรู้สึก (Sense) ผลการตอบจะได้คะแนนแน่นอน (ไม่มีการตอบผิด) เช่นเดียวกับการให้คะแนนของแบบวัด (Scale)

2. แบ่งตามลักษณะการตอบ แบ่งได้ 3 ประเภท คือ

2.1 แบบทดสอบภาคปฏิบัติ (Performance test) หมายถึง แบบทดสอบที่ให้นักเรียนลงมือปฏิบัติจริง เช่น การปรุงอาหาร การซ่อมแก้เครื่องยนต์ การฝีมือ ศิลปะ เป็นต้น



2.2 แบบทดสอบข้อเขียน (Paper-pencil test) หมายถึง แบบทดสอบที่ใช้การเขียนตอบ สามารถตั้งคำถามให้เขียนตอบจำนวนมาก ๆ หรือลึกซึ้งเพียงใดก็ได้ ผู้ตอบมีโอกาสเรียบเรียงเนื้อหา แลใช้ความสามารถทางสมองได้

2.3 แบบทดสอบปากเปล่า (Oral test) หมายถึง แบบทดสอบที่ใช้การพูดโต้ตอบแทนการเขียน เหมาะที่จะใช้กับผู้ที่ไม่ออกเขียนไม่ได้ หรือเมื่อต้องการให้ตอบอย่างฉับไว ลักษณะเช่นนี้ควรให้ตอบสั้น ๆ และมีข้อสอบไม่มากนัก เพราะจะเสียเวลามาก และต้องอาศัยการเรียบเรียงเนื้อหา

### 3. แบ่งตามเวลาที่กำหนดให้ตอบ แบ่งเป็น 2 ประเภท

3.1 แบบทดสอบที่จำกัดเวลาในการตอบ (Speed test) หมายถึง แบบทดสอบที่ใช้เวลาตอบน้อย แต่มีจำนวนข้อสอบมากและค่อนข้างง่าย ลักษณะเช่นนี้ต้องการจะทดสอบว่าในเวลาจำกัดใครจะมีความคล่องแคล่วรวดเร็วในการทำข้อสอบได้ถูกต้องมากกว่า การทดสอบในโรงเรียนส่วนใหญ่จะทำในลักษณะนี้

3.2 แบบทดสอบที่ไม่จำกัดเวลาในการตอบ (Power test) หมายถึง แบบทดสอบที่ใช้เวลาตอบมาก หรือไม่จำกัดเวลาในการตอบ แต่มีข้อสอบน้อย ลักษณะนี้ต้องการให้นักเรียนทุกคนได้แสดงศักยภาพของตนอย่างเต็มที่ มักจะเป็นข้อสอบที่ต้องแสดงความคิดเห็นหรือไหวไหวเคราะห์ บางครั้งจะให้เปิดหนังสือควบคู่ไปกับการสอบ หรือบางครั้งให้เอาไปตอบที่บ้าน

#### การใช้แบบทดสอบ

การใช้แบบทดสอบเพื่อการวัดผลการศึกษา นั้นควรคำนึงถึงหลักเกณฑ์และข้อจำกัดบางประการ ดังนี้

1. เลือกใช้ข้อสอบที่วัดคุณลักษณะที่ต้องการนั้นอย่างใกล้ชิดกับความเป็นจริงมากที่สุด
2. การใช้แบบทดสอบต้องให้เหมาะสมกับระดับกลุ่มของเด็ก
3. การใช้แบบทดสอบต้องให้เกิดความยุติธรรมมากที่สุดทั้งในด้านของข้อคำถามที่ใช้และวิธีดำเนินการสอบ
4. ควรสอบวัดหลาย ๆ ด้าน และใช้ข้อสอบหลาย ๆ อย่างประกอบกัน
5. แบบทดสอบที่ใช้ได้ผลตามเป้าหมายจะต้องเป็นแบบทดสอบที่มีคุณภาพดี
6. คะแนนจากการสอบเป็นเพียงตัวเลขในมาตราจัดอันดับ (Ordinal scale) จึงควรคำนึงถึงในการแปลคะแนนที่ได้จากการสอบ
7. การใช้แบบทดสอบควรพยายามใช้ผลการสอบที่ได้ให้กว้างขวางหลาย ๆ ด้าน

8. คะแนนที่ได้จากการสอบในแต่ละครั้ง ต้องถือว่าเป็นคะแนนที่เกิดจากความสามารถของแต่ละบุคคล ดังนั้นถ้ามีสิ่งใดที่จะทำให้การสอบได้ผลไม่ตรงกับความคิดข้างต้น ควรจะได้รับการแก้ไข

9. พึงระวังการสอบที่ทำให้เด็กได้เปรียบกัน เช่น การใช้แบบทดสอบชุดเดิม

10. การสอบแต่ละครั้งควรมีการตรวจให้คะแนนผลการสอบอย่างเป็นปรนัย

#### แบบทดสอบมาตรฐาน

เพ็ญศรี เศรษฐวงศ์ (2549, หน้า 161) ได้ให้ความหมายของ แบบทดสอบมาตรฐาน หมายถึง เครื่องมือวัดที่สามารถใช้เปรียบเทียบกับคนในกลุ่มอื่น ๆ ได้ โดยใช้คำถามชุดเดียวกัน สอบ/ วัดในสถานที่ต่างกัน แต่วิธีดำเนินการสอบ การให้คะแนน และเกณฑ์ในการแปลผล เป็นแบบเดียวกัน

พรพรรณ วันเอก (2552, หน้า 25) ได้ให้ความหมายของ แบบทดสอบมาตรฐาน หมายถึง แบบทดสอบวัดผลสัมฤทธิ์ของผู้เรียนทั่ว ๆ ไปหรือกลุ่มผู้วัดที่ต่างกลุ่มกัน และเป็นแบบทดสอบผ่านกระบวนการพัฒนาและปรับปรุงจนได้คุณภาพมาตรฐาน ระบุค่าความตรง ความเที่ยงของแบบทดสอบ ตลอดจนมีเกณฑ์ปกติ (Norm) สำหรับเป็นมาตรฐานการเปรียบเทียบ เพื่อแปลความหมายของคะแนน

จากความหมายที่กล่าวมาสรุปได้ว่า แบบสอบมาตรฐาน เป็นแบบทดสอบที่สร้างขึ้นมาโดยอาศัยการวางแผน การวิเคราะห์ปรับปรุงคุณภาพ มีการหาเกณฑ์ปกติเพื่อใช้เป็นหลักในการเปรียบเทียบผลสัมฤทธิ์ ส่วนใหญ่จะใช้ในการวัดผลสัมฤทธิ์ทางการเรียนแบบรวบยอด

แบบทดสอบมาตรฐานจะมีความเป็นมาตรฐานอยู่ 2 ประการ คือ

1. มาตรฐานในการดำเนินการสอบ เพื่อควบคุมตัวแปรที่จะมีผลกระทบต่อคะแนนของผู้สอบ ดังนั้นข้อสอบมาตรฐานจึงจำเป็นต้องมีคู่มือดำเนินการสอบไว้เป็นแนวปฏิบัติสำหรับผู้ใช้ข้อสอบ

2. มาตรฐานในการแปลความหมายคะแนน เพราะข้อสอบมาตรฐานมีเกณฑ์สำหรับการเปรียบเทียบคะแนนให้เป็นมาตรฐานเดียวกัน ซึ่งเรียกว่า เกณฑ์ปกติ (Norm) แบบทดสอบที่ครูสร้างขึ้นเอง มีข้อดีตรงที่ครูวัดได้ตรงจุดมุ่งหมายเพราะผู้สอนเป็นผู้ออกข้อสอบเอง แต่แบบทดสอบมาตรฐานมีข้อดีตรงที่คุณภาพของแบบทดสอบเป็นที่เชื่อถือได้ ทำให้สามารถนำผลไปเปรียบเทียบได้กว้างขวางกว่า

### ขั้นตอนของการสร้างข้อสอบมาตรฐาน

เพ็ญศรี เศรษฐวงศ์ (2549, หน้า 163) ได้กล่าวถึงขั้นตอนการสร้างแบบทดสอบไว้ ดังนี้

1. กำหนดรายละเอียดของแบบทดสอบ (Test specification) ก่อนที่จะสร้างแบบทดสอบ จะต้องมีการวางแผนการสร้างแบบทดสอบไว้ก่อนให้ละเอียดถี่ถ้วน แบ่งออกเป็นขั้นตอน ดังนี้

1.1 กำหนดวัตถุประสงค์ของแบบทดสอบ ทั้งวัตถุประสงค์ทั่วไป และวัตถุประสงค์เฉพาะ โดยต้องกำหนดสาขาเฉพาะที่จะวัด ใช้แบบทดสอบกับคนกลุ่มใด จะใช้คะแนนสอบอย่างไร ใช้เวลาสักเท่าไร มีการสร้างแบบทดสอบคู่ขนานด้วยหรือไม่ ขั้นนี้ถ้ามิใช่แบบทดสอบผลสัมฤทธิ์ธรรมดาทั่ว ๆ ไป เช่น เป็นแบบทดสอบวัดลักษณะทางจิตวิทยาต่าง ๆ ได้แก่ ความถนัด ความวิตกกังวล ความก้าวร้าว ฯลฯ จะต้องมีการนิยามตัวแปรที่จะวัดออกมาเป็นนิยามเชิงปฏิบัติการด้วยเพื่อบ่งบอกลักษณะของตัวแปรนั้น ซึ่งจะนำไปสู่การสร้างแบบทดสอบเพื่อวัดตัวแปรนั้นได้เที่ยงตรงดีขั้น ขั้นตอนในการกำหนดวัตถุประสงค์ของแบบทดสอบนี้เป็นขั้นตอนสำคัญ เพราะจะช่วยให้แบบทดสอบที่สร้างขึ้นมานั้นสามารถวัดในสิ่งที่ต้องการวัดได้ ขั้นตอนนี้ถ้าจะให้ได้ผลดีควรทำในรูปกรรมการและควรปรึกษาเจ้าหน้าที่เฉพาะ ผู้เชี่ยวชาญเฉพาะสาขาและควรศึกษาหลักสูตรอย่างละเอียด

1.2 สร้างตารางวิเคราะห์หลักสูตรหรือพิมพ์เขียวข้อสอบ มักทำ ในรูปของกรรมการ มีการวิเคราะห์เนื้อหาและพฤติกรรม แล้วจึงสังเคราะห์เป็นตารางวิเคราะห์หลักสูตรบ่งบอกว่า มีเนื้อหาอะไรบ้าง และในแต่ละเนื้อหา มีพฤติกรรมใดบ้าง และอย่างละเท่าไร แล้วจึงตัดแปลงให้เป็นพิมพ์เขียวข้อสอบบ่งบอกถึงโครงสร้างของแบบทดสอบทั้งในแง่เนื้อหาและพฤติกรรมหรือสมรรถภาพสมองที่จะวัดในแต่ละเนื้อหาเป็นการกำหนดเนื้อหาแบบทดสอบ

1.3 กำหนดเนื้อหาของข้อสอบ โดยแปลตารางวิเคราะห์หลักสูตรแต่ละช่องออกมาเป็นจุดประสงค์เชิงพฤติกรรม จนครบทุกช่องตามนี้หน้าของแต่ละช่องให้สัมพันธ์กับจำนวนข้อของแบบทดสอบ

1.4 เลือกแบบของข้อสอบที่เหมาะสม โดยพิจารณาถึงความเหมาะสมกับ วัตถุประสงค์ของแบบทดสอบ การตรวจให้คะแนน การพิมพ์ การดำเนินการสอบ ประหยัด วัดได้ครอบคลุม ให้ค่าความเชื่อมั่นและเที่ยงตรงสูง

1.5 กำหนดระดับความยากและการแจกแจงความยาก ข้อสอบที่เหมาะสมกับ วัตถุประสงค์ของแบบทดสอบ เหมาะกับค่าเฉลี่ยและความเบี่ยงเบนมาตรฐานที่ต้องการ

1.6 กำหนดจำนวนข้อของแบบทดสอบ โดยพิจารณาถึงนี้หน้าของตอนย่อย ในแบบทดสอบว่าตอนย่อยที่สามารถวัดได้ควรใช้อย่างน้อยกี่ข้อ, ทั้งหมดทุกตอนควรเป็นกี่ข้อ

คำนึงถึงความเชื่อมั่นของแบบทดสอบว่าควรจะมีมากข้อจึงจะมีความเชื่อมั่นสูง, คำนึงถึงช่วงเวลาที่ใช้ในการตอบแบบทดสอบฉบับนั้นว่าจะใช้เวลานานสักเท่าไร

1.7 กำหนดผู้เขียนและผู้ตรวจสอบบททวนข้อสอบ โดยต้องคำนึงถึงความสามารถ ประสบการณ์ความเชี่ยวชาญในการเขียนข้อสอบ มีความรู้ความเข้าใจในเรื่องการวัดผล และมีความรู้ความเข้าใจในเนื้อหาวิชานั้นเป็นอย่างดี

1.8 กำหนดตารางดำเนินงานแต่ละขั้นตอน เพื่อให้งานดำเนินไปตามแผนและสำเร็จ ล่วงไปได้ตั้งแต่เริ่มวางแผนจนกระทั่งจัดพิมพ์แบบทดสอบมาตรฐานเสร็จสิ้น

2. เขียนข้อสอบและตรวจทานข้อสอบ เมื่อเขียนข้อสอบเสร็จต้องมีการตรวจทาน ข้อสอบตามเกณฑ์ลักษณะข้อสอบที่ดี และวิเคราะห์เชิงตรรกวิทยาพร้อมทั้งแก้ไขปรับปรุงตรงจุด ที่เห็นว่าเป็นข้อบกพร่อง

3. จัดทำต้นฉบับแบบทดสอบ เริ่มด้วยการพิจารณาข้อสอบทั้งหมด แล้วประมาณ ความยากของข้อสอบแต่ละข้อ จัดเรียงเข้าเป็นฉบับเพื่อทดลองสอบ ขณะที่จัดทำต้นฉบับก็ต้อง พิจารณาความเหมาะสม ข้อบกพร่องในตัวด้วย เช่น ข้อสอบอาจพ้องกัน, แนะนำคำตอบกันรวมถึง พิจารณารูปแบบ และองค์ประกอบอื่น เช่น การวางหน้า, คำชี้แจง ฯลฯ หลังจากนั้นจึงนำไป จัดพิมพ์เพื่อทดลองสอบต่อไป

4. ทดลองสอบและปรับปรุงข้อสอบ ขั้นนี้เป็นการปรับปรุงคุณภาพข้อสอบเบื้องต้น โดยนำเอาแบบทดสอบไปทดลองสอบกับกลุ่มตัวอย่างของประชากรของแบบทดสอบนั้น นำผลการสอบมาวิเคราะห์รายข้อ ว่าดีเลว เด่นด้อยอย่างไรแล้วจึงปรับปรุงแก้ไข ตัดทอน เพิ่มเติม และวิเคราะห์ข้อสอบทั้งฉบับทั้งในแง่กายภาพและแง่สถิติ เช่น พิจารณาการเรียงอันดับข้อ พิจารณา ค่าสถิติพื้นฐาน เช่น ค่าเฉลี่ย ค่าความแปรปรวน แล้วจึงปรับปรุงการจัดฉบับใหม่ ใช้ค่าความยาก ที่วิเคราะห์ได้เป็นรากฐานในการจัดฉบับใหม่ นำไปทดลองสอบวิเคราะห์ ปรับปรุงเช่นนี้หลายครั้ง ซ้ำแล้วซ้ำอีกจนแน่ใจว่าข้อสอบแต่ละข้อมีคุณภาพสูงพอ

5. ตรวจสอบคุณภาพข้อสอบครั้งสุดท้าย โดยนำแบบทดสอบที่ทดลองแล้วปรับปรุง หลังจากนั้น ไปทดลองสอบกับกลุ่มตัวอย่างขนาดใหญ่พอควร เพื่อตรวจสอบคุณภาพข้อสอบ เป็นครั้งสุดท้ายแล้วหาค่าสถิติพื้นฐาน วิเคราะห์ค่าความยาก ค่าอำนาจจำแนก ค่าความเที่ยงตรง ค่าความเชื่อมั่นความคลาดเคลื่อนมาตรฐานและอื่น ๆ เมื่อแน่ใจว่าเป็นแบบทดสอบที่ดีพอจะเป็น แบบทดสอบมาตรฐานแล้วจึงนำไปสร้างเกณฑ์ปกติต่อไป

6. สร้างเกณฑ์ปกติ เมื่อได้แบบทดสอบที่มีคุณภาพสูงแล้วก็นำไปทดสอบกับ กลุ่มตัวอย่างที่เป็นตัวแทนประชากรแบบทดสอบนั้น เช่น ประชากรของแบบทดสอบเป็นนักเรียน ม.3 ทั่วประเทศ ก็ต้องเลือกกลุ่มตัวอย่างให้เป็นตัวแทนของนักเรียน ม.3 ทั่วประเทศ กลุ่มตัวอย่าง

นั้นต้องเป็นกลุ่มตัวอย่างใหญ่พอ เช่น อาจใช้ขนาดตั้งแต่ 5,000 คนขึ้นไป แต่ถ้าเป็นประชากรเล็กลงระดับท้องถิ่น เช่น ระดับจังหวัดใดจังหวัดหนึ่งก็อาจเป็นกลุ่มที่เล็กลงได้ การเลือกกลุ่มตัวอย่างอาจใช้กลุ่มตัวอย่างสุ่มหรือการสุ่มเป็นชั้นก็ได้ แต่นิยมใช้แบบหลังเพราะสามารถแบ่งชั้นได้หลายแบบทั้งแบบพื้นที่ แบบประเภท โรงเรียน ทำให้สามารถเลือกกลุ่มตัวอย่างที่เป็นตัวแทนกลุ่มประชากรได้ดี ขึ้นการออกเกณฑ์ปกตินี้ต้องดำเนินการสอบอย่างเคร่งครัดละเอียดถี่ถ้วนตามคำชี้แจงทุกประการ แล้วจึงนำคะแนนที่ได้มาแจกแจง แปลงเป็นคะแนนรูปอื่น นิยมใช้เปอร์เซ็นต์ไทล์หรือคะแนนมาตรฐานปกติแบบอื่น เช่น คะแนนทีปกติ (Normalized t-score) เป็นเกณฑ์ปกติ (Norms) ของแบบทดสอบฉบับนั้น อาจแสดงในรูปของตารางแปลงคะแนนหรือกราฟที่บ่งคะแนนดิบกับคะแนนที่แปลงแล้วในรูปอื่น ๆ

7. สร้างคู่มือการใช้แบบทดสอบ โดยบ่งรายละเอียดเกี่ยวกับวัตถุประสงค์ของแบบทดสอบบอกวิธีใช้ วิธีดำเนินการสอบ วิธีแปลความหมายคะแนน วิธีตรวจข้อสอบและวิธีการนำผลการสอบไปใช้

8. จัดพิมพ์แบบทดสอบจริง โดยการสำรวจตรวจสอบแต่ละข้ออีกครั้ง จัดรูปแบบให้เหมาะสมถูกต้องตามหลักวิชาการวัดผล แล้วจัดพิมพ์

จากที่ได้กล่าวมาสามารถสรุปหลักการสร้างข้อสอบ ดังนี้

1. หลักการสร้างข้อสอบนั้นผู้สร้างจะต้องยึดจุดประสงค์ในการสอนเป็นหลัก และจะต้องเขียนข้อสอบให้เป็นไปตามตารางวิเคราะห์หลักสูตรที่ได้กำหนดไว้ โดยจะต้องออกข้อสอบให้มีจำนวนข้อมากกว่าจำนวนข้อสอบในตารางวิเคราะห์หลักสูตร ทั้งนี้เพราะว่า เมื่อนำข้อสอบไปหาคุณภาพแล้วจะมีข้อสอบอยู่จำนวนหนึ่งที่ไม่ได้คุณภาพ ต้องตัดทิ้งไป

2. ข้อสอบมีหลายประเภททั้งนี้ขึ้นอยู่กับว่าจะยึดสิ่งใดเป็นเกณฑ์ในการแบ่ง

3. ข้อสอบที่ดีควรมีลักษณะ 5 ประการ เป็นอย่างน้อย คือ ความเที่ยงตรง ความเชื่อมั่น ความยาก อำนาจจำแนก และความมีประสิทธิภาพ

4. ข้อสอบอัตรานี้จะมีเฉพาะคำถามที่เปิดโอกาสให้ผู้ตอบได้แสดงออกโดยการเขียนที่แสดงถึงความรู้ความสามารถอย่างเต็มที่ ส่วนข้อสอบปรนัยจะมีทั้งคำถามและคำตอบมาให้ แล้วให้ผู้ตอบเลือกคำตอบที่ถูกต้อง ตามเงื่อนไขที่โจทย์กำหนด

5. ข้อสอบปรนัยสามารถแยกย่อยออกได้ 5 ประเภท คือ ข้อสอบแบบตอบสั้น ๆ แบบเติมคำ แบบจับคู่ แบบถูกผิดและแบบเลือกตอบ

**องค์ประกอบของแบบทดสอบมาตรฐาน**

ซิดเชื้อ กนกเพ็ชรรัตน์ (2550, หน้า 40) กล่าวว่าไว้ว่า องค์ประกอบ โดยทั่วไปของแบบทดสอบมาตรฐานมีอยู่ด้วยกัน 2 ส่วน

1. คู่มือการใช้แบบทดสอบโดยทั่วไปจะประกอบด้วยส่วนประกอบที่สำคัญ 6 ส่วน คือ
  - 1.1 จุดมุ่งหมายของแบบทดสอบ เป็นส่วนที่จะบอกถึงจุดประสงค์ในการสร้างแบบทดสอบใช้แบบทดสอบชนิดใด ใช้ทดสอบกับใคร วัตถุประสงค์กรมทางด้านใด
  - 1.2 วิธีสร้างแบบทดสอบ อธิบายถึงวิธีการสร้างโดยสรุปว่า ใช้ทดลองสอบ (Try out) กับกลุ่มตัวอย่างใด ทดสอบกี่ครั้ง มีค่าสถิติต่าง ๆ ในการทดสอบเท่าใด
  - 1.3 วิธีการดำเนินการสอบ จะต้องบอกวิธีการในการสอบโดยละเอียดว่าก่อนสอบระหว่างทำการสอบ และหลังสอบควรปฏิบัติอย่างไร เพื่อให้เป็นแบบอย่างเดียวกันในการสอบ
  - 1.4 วิธีการให้คะแนน ต้องระบุว่าถ้าตอบถูกในข้อใดได้คะแนนเท่าใด ถ้าตอบผิดได้คะแนนเท่าใด และต้องมีเฉลย (Key) ไว้ตรวจให้คะแนนด้วย
  - 1.5 วิธีแปลความหมายคะแนน ต้องระบุวิธีการตีความหมายของคะแนนจากผลการสอบ โดยสร้างเกณฑ์ปกติ หรือปกติวิสัย (Norms) ไว้สำหรับการเปรียบเทียบความสามารถของคนที่ทำแบบทดสอบ
  - 1.6 วิธีการนำผลการทดสอบไปใช้ ต้องระบุว่าจะนำผลการทดสอบไปใช้เพื่ออะไร เช่น การปรับปรุงการเรียนการสอน การตัดสินผลการเรียน การวิจัย เป็นต้น
2. ตัวแบบทดสอบ (Test) จะประกอบด้วย ชื่อแบบทดสอบ จุดมุ่งหมายในการทดสอบ คำชี้แจงต่าง ๆ ของแบบทดสอบโดยสรุปส่วนมากจะพิมพ์เป็นฉบับ และมีปกหน้าและปกหลัง โดยสรุปองค์ประกอบของแบบทดสอบมาตรฐานประกอบด้วย
  1. คู่มือการใช้แบบทดสอบโดยทั่วไปจะประกอบด้วยส่วนประกอบที่สำคัญ คือ จุดมุ่งหมายวิธีสร้าง วิธีดำเนินการสอบ วิธีการให้คะแนนวิธีแปลความหมายของคะแนน และการนำผลการทดสอบไปใช้
  2. ตัวแบบทดสอบจะประกอบด้วย ชื่อแบบทดสอบ จุดมุ่งหมายในการทดสอบคำชี้แจงต่าง ๆ ของแบบทดสอบ

### งานวิจัยที่เกี่ยวข้อง

ไชยยศ ไพวิทยศิริธรรม (2550, บทคัดย่อ) ได้ทำการวิจัยเรื่อง การพัฒนารูปแบบการประเมินหลักสูตร: การประยุกต์ใช้การประเมินอภิมาน เพื่อพัฒนารูปแบบการประเมินหลักสูตร โดยการประยุกต์ใช้วิธีการประเมินอภิมาน (Meta-evaluation) พบว่า 1) ผลการประเมินอภิมานงานวิจัยเกี่ยวกับการประเมินหลักสูตร ในภาพรวมอยู่ในระดับดีและเมื่อพิจารณาในแต่ละมาตรฐานพบว่า มาตรฐานด้านความเป็นไปได้ ด้านความเหมาะสม และด้านความถูกต้องอยู่ในระดับดี ยกเว้นมาตรฐานด้านการใช้ประโยชน์อยู่ในระดับพอใช้ ค่าคุณภาพผลการประเมินหลักสูตร

มีค่าเฉลี่ยเท่ากับ .0005 แสดงว่าผลการประเมินหลักสูตรอยู่ในระดับดี ผลการวิเคราะห์โมเดลเชิงเส้นตรงระดับลดหลั่น (HLM) พบว่า ตัวแปรปรับที่ส่งผลต่อค่าคุณภาพผลการประเมินหลักสูตรในระดับที่ 1 ได้แก่ ตัวแปรความเป็นมหาวิทยาลัยดั้งเดิมและตัวแปรการประเมินหลังการใช้หลักสูตร และตัวแปรปรับที่ส่งผลต่อค่าคุณภาพผลการประเมินหลักสูตรในระดับที่ 2 ได้แก่ ตัวแปรการประเมินหลังการใช้หลักสูตร ผลการวิเคราะห์เพื่อตรวจสอบความตรงของโมเดลเชิงสาเหตุของค่าคุณภาพผลการประเมินหลักสูตร พบว่า โมเดลมีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ โดยค่าไค-แอสควร์เท่ากับ 6.820 องศาอิสระเท่ากับ 3 ค่า p เท่ากับ .077 มีค่า GFI เท่ากับ .990 ค่า AGFI เท่ากับ .949 และค่า SRMR เท่ากับ .043 โมเดลสามารถร่วมกันอธิบายค่าความแปรปรวนของค่าคุณภาพผลการประเมินหลักสูตรได้ร้อยละ 14 ตัวแปรปีที่ทำวิจัยเสร็จ การประเมินหลังการใช้หลักสูตรบทบาทการประเมินแบบสรุปรวม และความเป็นมหาวิทยาลัยดั้งเดิมมีอิทธิพลต่อค่าคุณภาพผลการประเมินหลักสูตรอย่างมีนัยสำคัญ 2) รูปแบบการประเมินหลักสูตรที่พัฒนาขึ้นประกอบด้วย 5 ขั้นตอน คือ 1) การบริหารจัดการก่อนการประเมินหลักสูตร 2) การออกแบบการประเมินหลักสูตร 3) การดำเนินการประเมินหลักสูตร 4) การรายงานผลและเผยแพร่ผลการประเมินหลักสูตร 5) การประเมินผลการประเมินหลักสูตรผลการตรวจสอบความเหมาะสมของรูปแบบการประเมินหลักสูตร พบว่า รูปแบบการประเมินหลักสูตรมีความเหมาะสมโดยค่าเฉลี่ยเท่ากับ .78 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ .37

ทวีศักดิ์ อริยวัฒน์วงศ์ (2551, บทคัดย่อ) ได้ศึกษาเรื่องการประเมิน โครงการพัฒนาวิชนัยนักเรียนโรงเรียนวัดไร่จิงวิทยา ผลการวิจัยสรุปได้ ดังนี้ 1) ผลการวิเคราะห์สถานภาพครูผู้สอนและนักเรียน พบว่า ครูผู้สอนมีประสบการณ์ทางการสอน 21 ปี ขึ้นไป คิดเป็นร้อยละ 84.21 และการศึกษาระดับปริญญาตรีคิดเป็นร้อยละ 77.19 สอนช่วงชั้นที่ 3 คิดเป็นร้อยละ 73.68 ส่วนนักเรียนชั้นมัธยมศึกษาเป็นเพศชายคิดเป็นร้อยละ 47.34 เพศหญิง คิดเป็นร้อยละ 52.66 นักเรียนช่วงชั้นที่ 3 และ 4 เท่ากันคิดเป็นร้อยละ 50.00 รายได้ของผู้ปกครอง 10,001-20,000 บาท คิดเป็นร้อยละ 44.38 และผู้ปกครองมีอาชีพบริษัท/ ห้างร้านคิดเป็นร้อยละ 35.80 2) ผลการวิเคราะห์ข้อมูลการประเมิน โครงการพัฒนาวิชนัยนักเรียนของครูผู้สอน พบว่า ด้านบริบทโดยรวมอยู่ในระดับเหมาะสมมาก ด้านปัจจัยเบื้องต้นโดยรวมมีความเพียงพออยู่ในระดับมาก ด้านกระบวนการ โดยรวมมีการปฏิบัติงานอยู่ในระดับมาก ด้านผลผลิตโดยรวมมีความเป็นจริงอยู่ในระดับมาก และด้านผลกระทบครูมีความพึงพอใจโดยรวมอยู่ในระดับมาก 3) ผลการวิเคราะห์ข้อมูลด้านการพัฒนาวิชนัยนักเรียนโรงเรียนวัดไร่จิงวิทยา พบว่า ในภาพรวมอยู่ในระดับมาก

จุฑารัตน์ ทน้นไชย (2552, บทคัดย่อ) ได้ศึกษาเรื่อง การประยุกต์ใช้แบบจำลองของเคิร์กแพตทริกในการประเมินและติดตามผลโครงการฝึกอบรมการแก้ไขปัญหายาเสพติดใน

กลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด มีวัตถุประสงค์เพื่อศึกษาแบบจำลองของ เคิร์กแพตทริกในการประเมินและติดตามผลโครงการฝึกอบรม ข้อเด่นและข้อจำกัดของการใช้แบบ ในการประยุกต์ใช้ในการประเมินสัมฤทธิ์ผลของโครงการแก้ไขปัญหายาเสพติดในกลุ่มเยาวชน ผู้มีโอกาเข้าไปเกี่ยวข้องกับยาเสพติด ผลการวิจัย พบว่า ชั้นปฏิบัติการผู้เข้าร่วมโครงการ มีความพึงพอใจในการดำเนินการฝึกอบรมอยู่ในระดับดีมาก ในชั้นการเรียนรู้ผู้เข้าร่วมโครงการ มีความรู้เกี่ยวกับโรคเอดส์และยาเสพติดในระดับสูง ชั้นพฤติกรรมและชั้นผลลัพธ์ พบว่า ผู้เข้าร่วม โครงการมีการเปลี่ยนแปลงพฤติกรรมภายหลังจากการฝึกอบรมในบางพฤติกรรม และสามารถนำ ผลที่ได้รับจากการฝึกอบรมเรื่องยาเสพติดไปใช้ประโยชน์มากที่สุด ในส่วนของข้อเด่นในการใช้ แบบจำลองของเคิร์กแพตทริก คือ สามารถนำมาประยุกต์ใช้ได้กับเด็กและเยาวชนกลุ่มนี้ อย่างเหมาะสม ในส่วนของข้อจำกัดมี 2 ลักษณะ คือ ข้อจำกัดที่เกิดจากเครื่องมือ และจาก กระบวนการประยุกต์ใช้เครื่องมือในเรื่องสภาพแวดล้อมที่ไม่เอื้ออำนวย

สิรินธร สินจินดาวงศ์ (2552, บทคัดย่อ) ได้ทำการวิจัยเรื่อง การพัฒนาและประยุกต์ใช้ มาตรฐานการประเมินอภิมานสำหรับประเมินผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา ในประเทศไทย มีวัตถุประสงค์เพื่อ 1) พัฒนามาตรฐานการประเมินอภิมาน สำหรับประเมินรายงาน ผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาในประเทศไทย 2) ตรวจสอบ คุณภาพของมาตรฐานการประเมินอภิมาน ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพ ภายในของสถานศึกษา ระดับอุดมศึกษา 3) ประยุกต์ใช้มาตรฐานการประเมินอภิมานในการ ประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา ผลการวิจัย พบว่า 1) มาตรฐานการประเมินอภิมาน ประกอบด้วย 5 มาตรฐาน คือ ความตรงของการประเมิน อรรถประโยชน์ จริยธรรมในการประเมิน ความน่าเชื่อถือของนักประเมินและประเมินผลต้นทุน รวม 38 ตัวบ่งชี้ 2) คุณภาพของมาตรฐานการประเมินอภิมาน มีความตรงเชิงโครงสร้าง ด้วยเทคนิค กลุ่มรู้ชัด โดยสามารถจำแนกคุณภาพรายงานผลการประเมินภายใน ออกเป็นกลุ่มสูงและกลุ่มต่ำ และความเที่ยงแบบสัมประสิทธิ์การสุปร้องอิง เมื่อใช้นักประเมินอภิมาน 1, 2, 3 คน ประเมิน รายงานผลการประเมินคุณภาพภายใน จำนวน 3 ฉบับ มีค่าเท่ากับ 0.542, 0.689 และ 0.758 ตามลำดับ หากประเมินรายงานผลการประเมินคุณภาพภายใน 5 ฉบับ มีค่าเท่ากับ 0.663, 0.787 และ 0.839 ตามลำดับ ประเมินรายงานผลการประเมินคุณภาพภายใน 7 ฉบับ มีค่าเท่ากับ 0.734, 0.838 และ 0.880 ตามลำดับ และประเมินรายงานผลการประเมินคุณภาพภายใน 9 ฉบับ มีค่าเท่ากับ 0.780, 0.869 และ 0.904 ตามลำดับ 3) ผลการประเมินอภิมานของรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษาอยู่ในระดับดี โดยมาตรฐานที่มีผลการประเมินในระดับดี คือ ความน่าเชื่อถือ



ของนักประเมินอรรถประโยชน์ จริยธรรมในการประเมิน และความตรงของการประเมิน ส่วนมาตรฐานประสิทธิภาพต้นทุนมีผลประเมินอยู่ในระดับพอใช้

สันสนีย์ อุดมอ่าง (2554, บทคัดย่อ) ได้ประเมินโครงการ การฝึกอบรมเชิงปฏิบัติการ การผลิตทองม้วนเพื่อกระตุ้นเศรษฐกิจชุมชนตามแนวเศรษฐกิจพอเพียง พบว่า ผู้เข้าอบรม มีความพึงพอใจต่อการอบรมระดับมากถึงมากที่สุด โดยเฉพาะวิทยากรมีความรู้ความสามารถ ในการถ่ายทอดและปริมาณวัสดุอุปกรณ์ในการฝึก ผู้เข้าอบรมมีความรู้ความเข้าใจ ได้คะแนน หลังการอบรมเฉลี่ยสูงกว่าคะแนนก่อนการอบรมเฉลี่ย และร้อยละของค่าเฉลี่ยที่เปลี่ยนไปถึง ร้อยละ 49 ผู้เข้าอบรมได้นำความรู้จากการอบรมไปใช้ร้อยละ 0-30 ทั้งนี้เนื่องจากผู้เข้าอบรม เกือบทั้งหมด ไม่มีพิมพ์ทำทองม้วน ผลลัพธ์ที่เกิดกับชุมชน ชุมชนจะนำองค์ความรู้ไปใช้ประโยชน์ ในอนาคตเมื่ออำเภอหล่มเก่าเปิดตลาดร้อยปี จะใช้ทองม้วนเป็นผลิตภัณฑ์ของกลุ่มหล่มเก่าร่วมใจ สามัคคี

ชญัญญากัก วงศ์บา (2555) ได้ศึกษาเรื่อง ประเมินผลการฝึกอบรมตาม โครงการ CHAMPION โดยประยุกต์ใช้แนวทางการประเมินของเคิร์กแพตทริก ผลการวิจัย พบว่า ผู้เข้ารับ การฝึกอบรมส่วนใหญ่สังกัดหน่วยงานราชการ เป็นเพศหญิง ระดับการศึกษาสูงสุด จบปริญญาตรี อายุเฉลี่ย 41.53 ส่วนเบี่ยงเบนมาตรฐาน 8.31 การประเมินระดับที่ 1 พบว่า ส่วนใหญ่มีความคิดเห็น ความพึงพอใจต่อเนื้อหาหลักสูตรการฝึกอบรม วิทยากร/ เทคนิคการฝึกอบรม และปัจจัยเกื้อหนุน อยู่ในระดับมากถึงมากที่สุด และการประเมินระดับที่ 2 คะแนนเฉลี่ยก่อนการฝึกอบรม 15.71 ค่าส่วนเบี่ยงเบนมาตรฐาน 2.57 และคะแนนเฉลี่ยหลังการฝึกอบรม 25.02 ค่าส่วนเบี่ยงเบน มาตรฐาน 1.79 ความแตกต่างของคะแนนก่อนและหลังการฝึกอบรม ( $d$ ) เท่ากับ 9.31 (95% CI: 8.17 ถึง 10.45) และความรู้ความเข้าใจคะแนนเฉลี่ยก่อนและหลังการฝึกอบรมแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 (P-value < 0.001)

## บทที่ 3

### วิธีดำเนินการวิจัย

การวิจัยเรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียนในโรงเรียนสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเครื่องแพดทริกและการประเมินอภิมาน” มีวัตถุประสงค์เพื่อ เพื่อประเมินโครงการสร้างข้อสอบมาตรฐานวัดความสามารถในการคิดขั้นสูงของเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเครื่องแพดทริกและการประเมินอภิมาน โดยมีรายละเอียดเกี่ยวกับวิธีดำเนินการวิจัย ดังนี้

1. ประชากรและกลุ่มตัวอย่าง
2. เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล
3. วิธีการสร้างเครื่องมือ
4. การเก็บรวบรวมข้อมูล
5. การวิเคราะห์ข้อมูล
6. สถิติที่ใช้ในการวิจัย

#### ประชากรและกลุ่มตัวอย่าง

##### ประชากร

ประชากรในครั้งนี้เป็นบุคลากรครูในสังกัดสำนักงานการศึกษาเมืองพัทยาทั้ง 11 โรงเรียน จำนวน 609 คน

##### กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ คือ บุคลากรที่เข้าร่วมอบรมในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของเมืองพัทยา จำนวน 114 คน ประกอบด้วย หัวหน้าฝ่ายวิชาการจำนวน 1 คน นักวิชาการศึกษา/ เจ้าหน้าที่โครงการจำนวน 3 คน ผู้อำนวยการสถานศึกษา/ รองผู้อำนวยการสถานศึกษา (รับผิดชอบงานวิชาการ) จำนวน 10 คน ผู้แทนครูวัดผลทุกสถานศึกษา จำนวน 11 คน ครูผู้สอนกลุ่มสาระการเรียนรู้ทั้ง 8 กลุ่มสาระการเรียนรู้จำนวน 89 คน

## เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลการประเมินโครงการในครั้งนี้ ได้แก่

1. แบบประเมิน เป็นแบบประเมินที่ใช้วัดความรู้ความเข้าใจและความคิดเห็นของผู้เข้ารับการอบรม แบบประเมินจะเป็นเครื่องมือที่ผู้ศึกษาได้จัดเรียงอย่างเป็นระเบียบ เป็นชุดของคำถาม โดยจะมีทั้งข้อคำถามที่เป็นปลายปิดและปลายเปิด ซึ่งแนวของคำถามที่ใช้ในแบบประเมินจะเป็นไปตามรูปแบบการประเมินของเคิร์กแพททริกและการประเมินอภิมาน โดยแบ่งเป็น 3 ตอน ดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมิน

ตอนที่ 2 การประเมินด้านปฏิกิริยาความพึงพอใจในการเข้าร่วมโครงการ ด้านการเรียนรู้ ทักษะการปฏิบัติ ด้านพฤติกรรมการนำความรู้ที่ได้ไปปฏิบัติจริง และด้านผลลัพธ์ที่พึงประสงค์ของโครงการซึ่งบูรณาการตามรูปแบบการประเมินอภิมาน (4 มิติ)

โดยมีเกณฑ์การให้คะแนนความสำคัญของระดับความต้องการเป็นแบบมาตราส่วนประเมินค่า (Rating scale) มีตัวเลือก 5 ระดับ ดังนี้

คะแนน 5 หมายถึง ระดับอยู่ในเกณฑ์ เหมาะสมมากที่สุด หรือเห็นด้วยมากที่สุด

คะแนน 4 หมายถึง ระดับอยู่ในเกณฑ์ เหมาะสมมาก หรือเห็นด้วยมาก

คะแนน 3 หมายถึง ระดับอยู่ในเกณฑ์ เหมาะสมปานกลาง หรือเห็นด้วยปานกลาง

คะแนน 2 หมายถึง ระดับอยู่ในเกณฑ์ เหมาะสมน้อย หรือเห็นด้วยน้อย

คะแนน 1 หมายถึง ระดับอยู่ในเกณฑ์ เหมาะสมน้อยที่สุด หรือเห็นด้วยน้อยที่สุด

ตอนที่ 3 การประเมินอภิมาน (4 มิติ) และข้อคิดเห็น/ ข้อเสนอแนะ

2. การสัมภาษณ์ ใช้สำหรับเก็บข้อมูลในการประเมินและติดตามผลโครงการอบรม โดยจะเก็บจากผู้เกี่ยวข้องในโครงการ ได้แก่ เจ้าหน้าที่ของหน่วยงานที่รับผิดชอบ บุคลากรครูที่เข้าร่วมอบรม โดยสัมภาษณ์เพื่อสอบถามความเห็นซึ่งจะเก็บข้อมูลด้วยการสัมภาษณ์ทั้งอย่างเป็นทางการและไม่เป็นทางการ โดยข้อคำถามจะต้องการทราบ ดังนี้

2.1 ความเหมาะสมของโครงการ ความเหมาะสมของวัน-เวลา ความเหมาะสมของผู้เข้ารับการอบรม ความเหมาะสมของสถานที่ ตลอดจนควรมีการปรับปรุงอย่างไร

2.2 ประโยชน์ จากการเข้ารับการอบรมมีประโยชน์หรือไม่ และหลังจากได้อบรมแล้วจะนำความรู้ที่ได้ไปเผยแพร่ต่อเพื่อนร่วมงานอย่างไร

2.3 ปัญหาและอุปสรรค หลังจากที่ได้รับอบรมแล้ว พบปัญหาและอุปสรรคที่เกิดขึ้นอย่างไรเพื่อที่จะได้ทราบข้อเด่น-ข้อด้อยที่เกิดขึ้น และนำไปสู่การแก้ไขปรับปรุงในการอบรมครั้งต่อไป

## วิธีการสร้างเครื่องมือ

เครื่องมือที่ใช้ในการวิจัยครั้งนี้เป็นแบบประเมินสำหรับผู้เข้าร่วมการอบรมโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูง เป็นข้อมูลความคิดเห็นของผู้ตอบแบบสอบถามเกี่ยวกับโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูงโดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน ซึ่งประกอบด้วย

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมินเกี่ยวกับเพศ อายุ ประสบการณ์สอน ตำแหน่ง/ วิทยฐานะ ระดับชั้นที่สอน กลุ่มสาระการเรียนรู้ที่สอน คุณวุฒิสูงสุดทางการศึกษา วุฒิสายาอื่น ๆ

ตอนที่ 2 เป็นข้อมูลการประเมินโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูง ตามรูปแบบการประเมินของเคิร์กแพตทริก ร่วมกับการประเมินอภิมาน ซึ่งมีอยู่ 4 ด้าน ดังนี้

1. การประเมินด้านปฏิริยาความพึงพอใจร่วมกับความเหมาะสมของโครงการ
2. การประเมินด้านการเรียนรู้ทักษะการปฏิบัติร่วมกับความถูกต้องของเนื้อหา

ในการอบรม

3. การประเมินด้านพฤติกรรมการนำความรู้ที่ได้ไปปฏิบัติจริงร่วมกับความเป็นไปได้ หลังจากอบรมเสร็จแล้ว

4. การประเมินด้านผลลัพธ์ที่พึงประสงค์ของโครงการร่วมกับประโยชน์ในการอบรม

ตอนที่ 3 เป็นข้อมูลการประเมินโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูง ตามรูปแบบการประเมินอภิมาน (4 มิติ) และข้อคิดเห็น/ ข้อเสนอแนะเพื่อนำไปปรับปรุงโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูงให้มีประสิทธิภาพมากขึ้น

1. รรถประโยชน์ (Utility)
2. ความเป็นไปได้ (Feasibility)
3. ความเหมาะสม (Propriety)
4. ความถูกต้อง (Accuracy)

**ขั้นตอนในการสร้างเครื่องมือในการวิจัย**

แบบประเมินฉบับนี้ ผู้วิจัยได้สร้างขึ้นตามลำดับขั้นตอนในการสร้าง ดังนี้

1. ศึกษาแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้องในเรื่องการประเมิน โครงการสร้าง ข้อสอบมาตรฐานเพื่อวัดความรู้ ความสามารถในการคิดขั้นสูง จากนั้นมากำหนดกรอบแนวคิดที่ใช้ในการศึกษา กำหนดนิยามเพื่อใช้เป็นแนวทางในการสร้างแบบประเมิน โดยบูรณาการรูปแบบ การประเมิน โครงการแบบเคิร์กแพตทริกและการประเมินอภิमान

2. ศึกษาวิธีการสร้างแบบประเมิน และแบบสัมภาษณ์ จากเอกสารที่เกี่ยวข้องกับการ สร้างเครื่องมือในการวิจัย

3. ร่างคำถาม และจัดทำแบบประเมินฉบับร่าง สร้างแบบประเมินเกี่ยวกับการ ประเมิน โครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ ความสามารถในการคิดขั้นสูงและ ตรวจสอบแบบประเมินว่าครอบคลุมวัตถุประสงค์หรือไม่ จากนั้นนำไปให้อาจารย์ที่ปรึกษา ตรวจสอบและนำมาปรับปรุงตามคำแนะนำของอาจารย์ที่ปรึกษา

4. ให้ผู้ทรงคุณวุฒิพิจารณาแบบประเมิน นำแบบประเมินที่สร้างขึ้นที่ปรับปรุงตาม คำแนะนำของอาจารย์ที่ปรึกษาแล้วไปให้ผู้ทรงคุณวุฒิจำนวน 5 ท่าน ตรวจสอบและนำมาหาค่า ดัชนีความสอดคล้องของเครื่องมือ (Index of Item Objectives Congruence: IOC)

ให้คะแนน +1 เมื่อแน่ใจว่าข้อคำถามมีความสอดคล้องกับวัตถุประสงค์

ให้คะแนน 0 เมื่อไม่แน่ใจว่าข้อคำถามมีสอดคล้องกับวัตถุประสงค์หรือไม่

ให้คะแนน-1 เมื่อแน่ใจว่าข้อคำถามไม่สอดคล้องกับวัตถุประสงค์

5. หากคุณภาพแบบประเมิน คือ หากค่าอำนาจจำแนก และความเชื่อมั่น

6. นำไปใช้จริง

#### วิธีการหาคุณภาพแบบประเมิน

1. สถิติที่ใช้ในการตรวจสอบคุณภาพเครื่องมือ

1.1 การหาค่าความตรงตามเนื้อหา (Content validity) ของแบบประเมิน โดยการวัดดัชนีความสอดคล้อง

1.2 หาความเที่ยง (Reliability) ของแบบประเมิน โดยการหาค่าสัมประสิทธิ์แอลฟา ( $\alpha$ -Coefficient) ตามวิธีของครอนบาค (Cronbach)

#### การเก็บรวบรวมข้อมูล

ในการศึกษาวิจัยครั้งนี้ ได้ดำเนินการเก็บรวบรวมข้อมูลตามขั้นตอน ดังนี้

1. ประชุมหารือแนวทางในการดำเนินโครงการร่วมกับเจ้าหน้าที่ที่รับผิดชอบโครงการ

2. ใช้แบบประเมินที่ผ่านการตรวจสอบคุณภาพเครื่องมือแล้ว ไปเก็บรวบรวมข้อมูล โดยขอความอนุเคราะห์ในการแจกแบบประเมินผู้เข้าร่วมการอบรม โครงการสร้างข้อสอบ มาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูงตามระยะที่อบรม
3. ใช้แบบสัมภาษณ์ในการเก็บข้อมูลกับเจ้าหน้าที่ที่รับผิดชอบโครงการและบุคลากรครู ที่เข้าร่วมการอบรม เพื่อเป็นข้อมูลเพิ่มเติมในการประเมินและติดตามผลโครงการ
4. ผู้วิจัยรวบรวมแบบประเมินแล้วนำมาวิเคราะห์ทางสถิติตามกรอบการประเมิน

### การวิเคราะห์ข้อมูล

ผู้วิจัยได้ทำการวิเคราะห์ข้อมูลตามขั้นตอน ดังนี้

1. การตรวจสอบคุณภาพของเครื่องมือ ซึ่ง ได้แก่
  - 1.1 การหาค่าความตรงตามเนื้อหา (Content validity) ของแบบประเมิน โดยการวัดดัชนีความสอดคล้อง (0.8-1.0)
  - 1.2 หาค่าความเที่ยง (Reliability) ของแบบประเมิน โดยการหาค่าสัมประสิทธิ์แอลฟา ( $\alpha$ -Coefficient) ตามวิธีของครอนบาค (Cronbach) (0.957)
2. แบบประเมิน เมื่อได้รับแบบประเมินมาแล้วจึงใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป SPSS วิเคราะห์ข้อมูลตามขั้นตอน ดังนี้
  - 2.1 วิเคราะห์ค่าร้อยละ
  - 2.2 ค่าเฉลี่ย (Mean) และส่วนเบี่ยงเบนมาตรฐาน (SD)
  - 2.3 การเปรียบเทียบค่าเฉลี่ยโดยการทดสอบค่าที (Independent *t*-test)
  - 2.4 การเปรียบเทียบค่าเฉลี่ยโดยการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA)
  - 2.5 การศึกษาความสัมพันธ์โดยสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson product moment correlation)
  - 2.6 การวิเคราะห์เนื้อหา (Content analysis)

### สถิติที่ใช้ในการวิจัย

1. สถิติที่ใช้ในการตรวจสอบคุณภาพเครื่องมือ
  - 1.1 การหาค่าความตรงตามเนื้อหา (Content validity) ของแบบประเมิน โดยการวัดดัชนีความสอดคล้อง

1.2 หาความเที่ยง (Reliability) ของแบบประเมิน โดยการหาค่าสัมประสิทธิ์แอลฟา ( $\alpha$ -Coefficient) ตามวิธีของครอนบาค (Cronbach)

2. ค่าสถิติพื้นฐาน ได้แก่

2.1 ค่าคะแนนเฉลี่ย (Mean) โดยคำนวณจากสูตร

$$\bar{X} = \frac{\Sigma X}{N}$$

เมื่อ  $\bar{X}$  แทน คะแนนเฉลี่ย  
 $\Sigma X$  แทน ผลรวมของคะแนนทั้งหมด  
 $N$  แทน จำนวนผู้ตอบแบบประเมิน

2.2 ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) โดยคำนวณจากสูตร

$$SD = \sqrt{\frac{n\Sigma x^2 - (\Sigma x)^2}{n(n-1)}}$$

เมื่อ  $SD$  แทน ค่าส่วนเบี่ยงเบนมาตรฐาน  
 $X$  แทน คะแนนแต่ละคน  
 $\Sigma x$  แทน ผลรวมของคะแนนบุคลากรทั้งหมด  
 $x^2$  แทน กำลังสองของคะแนนแต่ละคน  
 $\Sigma x^2$  แทน ผลรวมของกำลังสองของคะแนนทั้งหมด  
 $n$  แทน จำนวนบุคลากรในโรงเรียน

3. สถิติที่ใช้ในการประเมินตัวบ่งชี้โครงการตามกรอบแนวคิดในการประเมิน

ตารางที่ 3 กรอบแนวคิดในการประเมิน (Evaluation framework)

วัตถุประสงค์	ตัวบ่งชี้	วิธีการ/ เครื่องมือ	แหล่งข้อมูล	สถิติ	เกณฑ์
เพื่อประเมินปฏิกิริยาตอบสนอง (Reaction evaluation)	1. สถานที่ 2. วิทยากร 3. เอกสาร 4. สิ่งอำนวยความสะดวก 5. โสตทัศนูปกรณ์	แบบประเมิน การสัมภาษณ์	เจ้าหน้าที่และ ผู้เข้าร่วมโครงการ	1. ค่าเฉลี่ย 2. ค่าส่วนเบี่ยงเบนมาตรฐาน 3. วิเคราะห์เนื้อหา	มาตรฐานค่า (Rating scale) 3.5 ขึ้นไป
เพื่อประเมินการเรียนรู้ (Learning evaluation)	1. ด้านความรู้ 2. ด้านความคิด 3. ด้านทักษะ/ การปฏิบัติ 4. ทักษะ/ ค่านิยม	แบบประเมิน	ผู้เข้าร่วมโครงการ	1. ร้อยละ 2. ค่าเฉลี่ย 3. ค่าส่วนเบี่ยงเบนมาตรฐาน	มาตรฐานค่า (Rating scale) 3.5 ขึ้นไป
เพื่อประเมินพฤติกรรม (Behavior evaluation)	1. การนำความรู้ไปประยุกต์ใช้ 2. การพัฒนาความรู้เพิ่มเติม 3. ความสามารถในการแก้ปัญหา 4. การนำความรู้ไปเผยแพร่ แลกเปลี่ยนเรียนรู้กับผู้ร่วมงาน	แบบประเมิน	ผู้เข้าร่วมโครงการ	1. ค่าเฉลี่ย 2. ค่าส่วนเบี่ยงเบนมาตรฐาน	มาตรฐานค่า (Rating scale) 3.5 ขึ้นไป
เพื่อประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation)	1. การลดลงของปัจจัยเสี่ยง 2. การเพิ่มประสิทธิภาพการทำงาน 3. การยอมรับขององค์กร	แบบประเมิน การสัมภาษณ์	ผู้เข้าร่วมโครงการ	1. ค่าเฉลี่ย 2. ค่าส่วนเบี่ยงเบนมาตรฐาน 3. วิเคราะห์เนื้อหา	มาตรฐานค่า (Rating scale) 3.5 ขึ้นไป



## บทที่ 4

### ผลการวิเคราะห์ข้อมูล

การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริก และการประเมินอภิมาน มีวัตถุประสงค์เพื่อประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน โดยมีวัตถุประสงค์เฉพาะ เพื่อประเมินปฏิกิริยาตอบสนอง (Reaction evaluation) เพื่อประเมินการเรียนรู้ (Learning evaluation) เพื่อประเมินพฤติกรรม (Behavior evaluation) และเพื่อประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน

ผลการศึกษาแบ่งเป็น 4 ตอน ดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมิน

ตอนที่ 2 ผลการวิเคราะห์ความพึงพอใจของผู้เข้าร่วมโครงการสร้าง

ข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา

ตอนที่ 3 การวิเคราะห์เปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน

ตอนที่ 4 การวิเคราะห์ความสัมพันธ์

## ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมิน

ตารางที่ 4 จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา ที่ตอบแบบประเมินจำแนกตาม เพศ อายุ และประสบการณ์การสอน

ข้อมูล	จำนวน	ร้อยละ
<b>เพศ</b>		
ชาย	12	10.50
หญิง	102	89.50
รวม	114	100.00
<b>อายุ</b>		
20-30 ปี	5	4.40
31-40 ปี	58	50.90
41-50 ปี	29	25.40
51-60 ปี	22	19.30
รวม	114	100.00
<b>ประสบการณ์การสอน</b>		
0-10 ปี	56	49.10
11-20 ปี	39	34.20
21-30 ปี	17	14.90
31-40 ปี	2	1.80
รวม	114	100.00

จากตารางที่ 4 พบว่า ข้อมูลส่วนบุคคลเมื่อจำแนกตามเพศ อายุ และประสบการณ์การสอนของผู้เข้าอบรมโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา จำแนกตามเพศปรากฏว่าส่วนใหญ่ เป็นเพศหญิง คิดเป็นร้อยละ 89.50 และเป็นเพศชาย คิดเป็นร้อยละ 10.50 เมื่อพิจารณาตามอายุของผู้เข้าอบรม ส่วนใหญ่มีอายุอยู่ระหว่าง 31-40 ปี คิดเป็นร้อยละ 50.90 รองลงมาอยู่ระหว่าง 41-50 ปี คิดเป็นร้อยละ 25.40 และอายุระหว่าง 20-30 ปีน้อยสุด คิดเป็นร้อยละ 4.40 และเมื่อพิจารณาตามประสบการณ์ที่สอน

พบว่า ส่วนใหญ่มีประสบการณ์ในการสอน 1-10 ปี คิดเป็นร้อยละ 49.10 รองลงมา 11-20 ปี คิดเป็นร้อยละ 34.20 และ 31-40 ปี น้อยที่สุด คิดเป็นร้อยละ 1.80

ตารางที่ 5 จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐาน วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ที่ออกแบบประเมินจำแนกตามตำแหน่งวิทยฐานะ และระดับชั้นที่สอน

ข้อมูล	จำนวน	ร้อยละ
ตำแหน่งวิทยฐานะ		
ครูผู้ช่วย	5	4.40
ครู	24	21.10
ครูชำนาญการ	52	45.60
ครูชำนาญการพิเศษ/ ครูเชี่ยวชาญ	6	5.30
อื่น ๆ (เจ้าหน้าที่, ผู้ช่วยครู)	27	23.70
รวม	114	100.00
ระดับชั้นที่สอน		
ประถมต้น	16	14.00
ประถมปลาย	36	31.60
มัธยมศึกษาตอนต้น	35	30.70
มัธยมศึกษาตอนปลาย	13	11.40
อื่น ๆ (เจ้าหน้าที่, ผู้บริหาร)	14	12.30
รวม	114	100.00

จากตารางที่ 5 พบว่า ข้อมูลส่วนบุคคลเมื่อจำแนกตามตำแหน่งวิทยฐานะ และระดับชั้นที่สอนของผู้เข้าร่วมโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ส่วนใหญ่ดำรงตำแหน่งครูชำนาญการ คิดเป็นร้อยละ 45.60 รองลงมาครู คิดเป็นร้อยละร้อยละ 21.10 และครูชำนาญการพิเศษ/ ครูเชี่ยวชาญน้อยที่สุด คิดเป็นร้อยละ 5.30 และเมื่อพิจารณาตามระดับชั้นที่สอน พบว่า ส่วนใหญ่สอนในระดับประถมปลาย คิดเป็นร้อยละ 31.60 รองลงมาสอนในระดับมัธยมศึกษาตอนต้น คิดเป็นร้อยละ 30.70 และสอนในระดับมัธยมศึกษาตอนปลายน้อยที่สุด คิดเป็นร้อยละ 11.40

ตารางที่ 6 จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา ที่ตอบแบบประเมินจำแนกตามกลุ่มสาระการเรียนรู้

ข้อมูล	จำนวน	ร้อยละ
กลุ่มสาระการเรียนรู้		
ภาษาไทย	14	12.30
คณิตศาสตร์	13	11.40
วิทยาศาสตร์	12	10.50
สังคมศึกษา ศาสนา และวัฒนธรรม	12	10.50
สุขศึกษาและพลศึกษา	12	10.50
ศิลปะ	12	10.50
การงานอาชีพและเทคโนโลยี	13	11.40
ภาษาต่างประเทศ	12	10.50
อื่น ๆ (เจ้าหน้าที่, ผู้บริหาร)	14	12.30
รวม	114	100.00

จากตารางที่ 6 พบว่า ข้อมูลส่วนบุคคลเมื่อจำแนกตามกลุ่มสาระการเรียนรู้ที่สอนของผู้เข้าอบรมโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา เมื่อพิจารณาตามกลุ่มสาระการเรียนรู้ที่สอน ส่วนใหญ่สอนในกลุ่มสาระการเรียนรู้ภาษาไทยและอื่น ๆ คือ เจ้าหน้าที่และผู้บริหาร คิดเป็นร้อยละ 12.30 รองลงมาสอนคณิตศาสตร์และการงานอาชีพและเทคโนโลยี คิดเป็นร้อยละ 11.40 และสอนวิทยาศาสตร์ สังคมศึกษา ศาสนาและวัฒนธรรม สุขศึกษา ศิลปะ และภาษาต่างประเทศ คิดเป็นร้อยละ 10.50

ตารางที่ 7 จำนวนและค่าร้อยละของผู้เข้าอบรมในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา  
ที่ตอบแบบประเมินจำแนกตามวุฒิสูงสุด และจำแนกตามสาขา

ข้อมูล	จำนวน	ร้อยละ
วุฒิการศึกษาสูงสุด		
ปริญญาตรี	68	59.60
ปริญญาโท	46	40.40
รวม	114	100.00
วุฒิจำแนกตามสาขา		
วุฒิทางการศึกษา	103	90.40
วุฒิสาขาคืออื่น ๆ	11	9.60
รวม	114	100.00

จากตารางที่ 7 พบว่า เมื่อจำแนกตามวุฒิการศึกษาสูงสุด และวุฒิจำแนกตามสาขาของผู้เข้าอบรมโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา เมื่อพิจารณาตามวุฒิการศึกษาสูงสุดส่วนใหญ่มีวุฒิการศึกษาปริญญาตรี คิดเป็นร้อยละ 59.60 และมีวุฒิการศึกษาปริญญาโท คิดเป็นร้อยละ 40.40 และเมื่อพิจารณาตามวุฒิโดยจำแนกตามสาขา พบว่า มีวุฒิทางการศึกษา คิดเป็นร้อยละ 90.40 และวุฒิสาขาคืออื่น ๆ ร้อยละ 9.60

ตอนที่ 2 ผลการวิเคราะห์ความพึงพอใจของผู้เข้าร่วมโครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัด เมืองพัทยา

ตารางที่ 8 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา  
ด้านปฏิริยาตอบสนอง

ข้อที่	รายการประเมิน	$\bar{X}$	$SD$	ผลประเมิน
1	ความเหมาะสมของสถานที่ในการจัดโครงการ	3.61	0.65	มาก
2	ความเหมาะสมของวัน-เวลาในการจัดโครงการ	3.30	0.91	ปานกลาง
3	ความเหมาะสมของวิทยากร	3.54	0.78	มาก
4	ความเหมาะสมของเอกสารประกอบการ บรรยาย	3.47	0.73	ปานกลาง
5	ความเหมาะสมหลักสูตรในแต่ละกิจกรรมกับ จำนวนเวลาในการ ฝึกอบรม	3.68	0.83	มาก
6	ความเหมาะสมของสิ่งอำนวยความสะดวก	3.81	0.51	มาก
7	ความเหมาะสมเพียงพอของสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในการฝึกอบรม	4.08	0.63	มาก
8	ความเหมาะสมของชนิดวัสดุทัศนูปกรณ์	3.82	0.51	มาก
9	ความเหมาะสมของความพร้อมอุปกรณ์ ทัศนูปกรณ์	3.99	0.52	มาก
10	ความเหมาะสมการลำดับกิจกรรมในการ ฝึกอบรม	3.70	0.73	มาก
รวมเฉลี่ย		3.70	0.47	มาก

จากตารางที่ 8 พบว่า ผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ด้านปฏิริยาตอบสนอง  
โดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 3.70, SD = 0.47$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า  
ความเหมาะสมเพียงพอของสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในการฝึกอบรมมีคะแนนเฉลี่ยสูง  
ที่สุด ( $\bar{X} = 4.08, SD = 0.63$ ) รองลงมา ได้แก่ ความเหมาะสมของความพร้อมอุปกรณ์

โตตัทสนูปกรณ์ ( $\bar{X} = 3.99, SD = 0.52$ ) ส่วนความเหมาะสมของวัน-เวลาในการจัดโครงการ มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.30, SD = 0.91$ )

สรุปประเด็นความคิดเห็นในด้านความเหมาะสมของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำนวน 15 คน 18 ความคิดเห็น

ตารางที่ 9 สรุปความคิดเห็นในด้านความเหมาะสมของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา

สรุปจากความคิดเห็น	ตัวอย่างความคิดเห็น
1. ระยะเวลาในการอบรม	<ol style="list-style-type: none"> <li>ระยะเวลาในการอบรมมีจำนวนระยะเวลาในการอบรมมีหลายช่วงเวลาทำให้กระทบกับการสอนของครู</li> <li>การจัดทำโครงการมีระยะเวลาทับซ้อนกับโครงการอื่น และยาวนานเกินไป ผู้อบรมบางท่านต้องปฏิบัติหน้าที่หลายอย่างในเวลาเดียวกัน</li> <li>เป็นการเข้าร่วมโครงการที่ใช้ระยะเวลานานเพิ่มมากขึ้นในการเข้าร่วมอบรมในโครงการนี้พอสมควรแต่ก็ได้ความรู้</li> </ol>
2. สถานที่	<ol style="list-style-type: none"> <li>สถานที่ในการอบรมดีมาก</li> <li>สถานที่ในการอบรมดี สะดวกต่อผู้เข้าร่วมการอบรมเพราะติดถนนสุขุมวิท</li> <li>บรรยากาศในการอบรมเป็นกันเองแต่บางครั้งเครียดเกินไป</li> <li>จัดตรงกับวันที่มีนักเรียนมาเรียนจึงทำให้บางครั้งมีเสียงมารบกวนเวลาทำกิจกรรม</li> </ol>
3. สิ่งอำนวยความสะดวก	<ol style="list-style-type: none"> <li>ดี มีสิ่งอำนวยความสะดวกตามที่ต้องการ</li> <li>เหมาะสม</li> </ol>

จากตารางที่ 9 พบว่า ความคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความเหมาะสมของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ระยะเวลาในการอบรมมีผู้ให้ความคิดเห็นทั้งในด้านบวกและด้านลบ ด้านบวก ดังนี้ “ดิฉันคิดว่าระยะเวลาในการอบรมครั้งนี้เหมาะสมดี” ในด้าน

ลบ “ดิฉันคิดว่าการอบรมมีจำนวนระยะหลายช่วงเวลาทำให้กระทบกับการสอนของครู, การจัดทำโครงการมีระยะเวลาอันยาวนานและทับซ้อนกับโครงการอื่น และยาวนานเกินไป ผู้อบรมบางท่านต้องปฏิบัติหน้าที่หลายอย่างในเวลาเดียวกัน” กลุ่มที่ 2 สถานที่ในการจัดการอบรมมีผู้ให้ความคิดเห็นทั้งด้านบวกและด้านลบ ในด้านบวก ดังนี้ “ดิฉันคิดว่าสถานที่ในการอบรมดี สะดวกต่อผู้เข้าร่วมการอบรมเพราะติดถนนสุขุมวิท การเดินทางสะดวก, บรรยากาศในการอบรมเป็นกันเอง” ในด้านลบ “ดิฉันคิดว่าบางครั้งเครียดเกินไปและจัดตรงกับวันที่มีนักเรียนมาเรียนจึงทำให้บางครั้งมีเสียงมารบกวนเวลาทำกิจกรรม” กลุ่มสุดท้ายสิ่งอำนวยความสะดวกมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าดี มีสิ่งอำนวยความสะดวกเพียงพอต่อความต้องการและเหมาะสม”

ตารางที่ 10 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา ด้านการเรียนรู้

ข้อที่	รายการประเมิน	$\bar{X}$	SD	ผลประเมิน
1	ความรู้ ความเข้าใจในเรื่องนี้หลังการอบรม	4.22	0.56	มาก
2	สามารถจัดระบบความคิด	3.96	0.60	มาก
3	สามารถประมวลความคิดสู่การพัฒนางาน อย่างเป็นระบบ	4.11	0.60	มาก
4	เกิดทักษะและประสบการณ์ใหม่ ๆ จาก การฝึกอบรม	3.91	0.65	มาก
5	สามารถปฏิบัติการออกแบบทดสอบวัดความรู้ ความสามารถขั้นสูงของนักเรียนได้	4.11	0.59	มาก
6	ประโยชน์ที่ท่านได้รับจากโครงการ	4.04	0.65	มาก
7	สิ่งที่ท่านได้รับจากโครงการนี้ตรงตามความ คาดหวัง	3.96	0.77	มาก
8	เนื้อหาที่บรรยายถูกต้องตรงกับหัวข้อที่กำหนด	4.07	0.53	มาก
	รวมเฉลี่ย	4.05	0.50	มาก

จากตารางที่ 10 พบว่า ผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา ด้านการเรียนรู้โดยภาพ



รวมอยู่ในระดับมาก ( $\bar{X} = 4.05, SD = 0.50$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า ความรู้ความเข้าใจในเรื่องนี้หลังการอบรมมีคะแนนเฉลี่ยสูงสุดที่สุด ( $\bar{X} = 4.22, SD = 0.56$ ) รองลงมา ได้แก่ สามารถประมวลความคิดสู่การพัฒนางานอย่างเป็นระบบกับข้อ 6 สามารถปฏิบัติการออกแบบทดสอบวัดความรู้ความสามารถขั้นสูงของนักเรียนได้ ( $\bar{X} = 4.11, SD = 0.59$ ) ส่วนความเกิดทักษะและประสบการณ์ใหม่ ๆ จากการฝึกอบรม มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.91, SD = 0.65$ )

สรุปประเด็นความคิดเห็นในด้านความถูกต้องด้านเนื้อหาโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำนวน 10 คน 14 ความคิดเห็น โดยผู้เข้าร่วมการอบรมได้แสดงความคิดเห็นไว้ ดังนี้

ตารางที่ 11 สรุปความคิดเห็นในด้านเนื้อหาของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา

สรุปจากความคิดเห็น	ตัวอย่างความคิดเห็น
1. ด้านเนื้อหา	<ol style="list-style-type: none"> <li>1. ความถูกต้องด้านเนื้อหาอยู่ในระดับดีมาก การให้ความรู้ในรายละเอียดของการออกข้อสอบจากคณะวิทยากรให้คำชี้แนะ แก้ไขและปรับปรุง</li> <li>2. ได้นำสาระแกนกลางปี พ.ศ. 2551 มาออกข้อสอบในแต่ละตัวชี้วัด</li> <li>3. เป็นโครงการที่ดี แต่มีเนื้อหาเยอะเข้าใจยาก</li> </ol>
2. ความรู้จากวิทยากร	<ol style="list-style-type: none"> <li>1. ให้คำอธิบายเข้าใจง่าย</li> <li>2. วิทยากรประจำกลุ่มให้การช่วยเหลือดี</li> </ol>
3. การนำเสนอข้อมูล	<ol style="list-style-type: none"> <li>1. ควรมีการจัดลำดับการนำเสนอข้อมูลต่อผู้เข้ารับการอบรม</li> <li>2. มีการนำเสนอล่าช้าผู้ปฏิบัติเกิดการสับสนและต้องแก้ไขข้อมูลใหม่ทำให้การทำงานล่าช้า</li> </ol>

จากตารางที่ 11 พบว่า ความคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความถูกต้องด้านเนื้อหาของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านเนื้อหาที่มีผู้ให้ความเห็นทั้งในด้านบวกและด้านลบ ด้านบวก ดังนี้ “คิดว่าคุณค่าความถูกต้องด้านเนื้อหาอยู่ในระดับดีมาก การให้ความรู้ในรายละเอียดของการออกข้อสอบจากคณะวิทยากรให้คำชี้แนะ แก้ไขและปรับปรุง,

มีการนำเสนอระแคนกลางปี 2551 มาออกข้อสอบในแต่ละตัวชี้วัด” ด้านลบ ดังนี้ “ดิฉันคิดว่าเป็นโครงการที่ดี แต่มีเนื้อหาเยอะเข้าใจยาก, เนื้อหาในการบรรยายเยอะ” กลุ่มที่ 2 ด้านความรู้จากวิทยากรมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าวิทยากรอธิบายเข้าใจง่ายและวิทยากรประจำกลุ่มให้การช่วยเหลือดี ” และกลุ่มที่ 3 ด้านการนำเสนอข้อมูลมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าควรมีการจัดลำดับการนำเสนอข้อมูลต่อผู้เข้ารับการอบรม, มีการนำเสนอล่าช้าผู้ปฏิบัติเกิดการสับสน และต้องแก้ไขข้อมูลใหม่ทำให้การทำงานล่าช้า

ตารางที่ 12 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ด้านพฤติกรรม

ข้อที่	รายการประเมิน	$\bar{X}$	SD	ผลประเมิน
1	สามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการปฏิบัติงานได้	4.18	0.65	มาก
2	สามารถนำความรู้ไปเผยแพร่/ ถ่ายทอดแก่เพื่อนร่วมงานได้	3.91	0.84	มาก
3	สามารถให้คำปรึกษาแก่เพื่อนร่วมงานได้	4.06	0.72	มาก
4	มีความมั่นใจและสามารถนำความรู้ที่ได้รับไปใช้ได้	3.94	0.82	มาก
5	สามารถแลกเปลี่ยนเรียนรู้กับเพื่อนร่วมงานได้	3.96	0.72	มาก
รวมเฉลี่ย		4.01	0.64	มาก

จากตารางที่ 12 พบว่า ผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ด้านพฤติกรรม โดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 4.01, SD = 0.64$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่าสามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการปฏิบัติงานได้มีคะแนนเฉลี่ยสูงที่สุด ( $\bar{X} = 4.18, SD = 0.65$ ) รองลงมา ได้แก่ สามารถให้คำปรึกษาแก่เพื่อนร่วมงานได้ ( $\bar{X} = 4.06, SD = 0.72$ ) ส่วนสามารถนำความรู้ไปเผยแพร่/ ถ่ายทอดแก่เพื่อนร่วมงานได้ มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.91, SD = 0.84$ )

สรุปประเด็นความคิดเห็นในด้านความเป็นไปได้จากโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำนวน 18 คน  
18 ความคิดเห็น

ตารางที่ 13 สรุปความคิดเห็นในด้านความเป็นไปได้หลังเข้ารับการอบรมโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา

สรุปจากความคิดเห็น	ตัวอย่างความคิดเห็น
1. มีข้อสอบที่มีมาตรฐาน	<ol style="list-style-type: none"> <li>1. ทำให้ได้ข้อสอบที่มีคุณภาพมากขึ้นในแต่ละรายวิชาที่สามารถจะวัดผู้เรียนได้ตรงตามมาตรฐานตัวชี้วัด</li> <li>2. มีการสร้างข้อสอบที่เป็นระบบและได้มาตรฐานตรงตามตัวชี้วัด</li> <li>3. นำข้อสอบที่สร้างไปใช้ทดสอบกับนักเรียนได้จริง</li> </ol>
2. พัฒนาตนเอง	<ol style="list-style-type: none"> <li>1. พัฒนาการจัดกิจกรรมการเรียนรู้ การสอนของครูให้สอดคล้องตรงตามแผนจัดการเรียนรู้ของมาตรฐานและตัวชี้วัด</li> <li>2. มีความเข้าใจในการทำข้อสอบเพิ่มมากขึ้นและได้มีเกณฑ์มาตรฐานในการทำข้อสอบมากเท่าที่ควรพร้อมทั้งยังสามารถนำไปปรับปรุงข้อสอบของตนเองได้</li> <li>3. นำความรู้ที่ได้รับมาปรับปรุงข้อสอบที่ใช้วัดกับเด็กนักเรียนและพัฒนาการเรียนการสอนได้</li> </ol>
3. เผยแพร่ความรู้แก่เพื่อนร่วมงาน	<ol style="list-style-type: none"> <li>1. เรารู้ขั้นตอน วิธีการที่จะจัดสร้างข้อสอบสามารถนำไปเผยแพร่กับเพื่อนร่วมงานได้</li> <li>2. สามารถให้คำปรึกษาแก่เพื่อนร่วมงานในด้านการวิเคราะห์ตัวชี้วัดเพื่อจัดทำแผนผังข้อสอบ (Test blueprint) และการวิเคราะห์ข้อสอบรายข้อได้</li> </ol>

จากตารางที่ 13 พบว่า ความคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความเป็นไปได้ของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านมีข้อสอบที่มีมาตรฐานมีผู้ให้

ความคิดเห็น ดังนี้ “ดิฉันคิดว่าจากการที่ได้ปฏิบัติกิจกรรมทำให้ได้ข้อสอบที่มีคุณภาพมากขึ้น ในแต่ละรายวิชาที่สามารถจะวัดผู้เรียนได้ตรงตามมาตรฐานตัวชี้วัด, มีการสร้างข้อสอบที่เป็นระบบ และได้มาตรฐานตรงตามตัวชี้วัด, สามารถนำข้อสอบที่สร้างไปใช้ทดสอบกับนักเรียนได้จริง”

กลุ่มที่ 2 ด้านการพัฒนาตนเองมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าเข้าใจในการทำข้อสอบ เพิ่มมากขึ้นและได้มีเกณฑ์มาตรฐานในการทำข้อสอบมากเท่าที่ควรพร้อมทั้งยังสามารถนำไปปรับปรุงข้อสอบของตนเองได้, มีการพัฒนาการจัดกิจกรรมการเรียนรู้ การสอนของครูให้สอดคล้องตรงตามแผนจัดการเรียนรู้ของมาตรฐานและตัวชี้วัดและนำความรู้ที่ได้รับมาปรับปรุงข้อสอบที่ใช้วัดกับเด็กนักเรียนและพัฒนาการเรียนการสอนได้”

กลุ่มที่ 3 ด้านการเผยแพร่ความรู้แก่เพื่อนร่วมงานมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าสามารถให้คำปรึกษาแก่เพื่อนร่วมงานในด้านการวิเคราะห์ตัวชี้วัดเพื่อจัดทำแผนผังข้อสอบ (Test blueprint) และการวิเคราะห์ข้อสอบรายข้อได้, รู้ขั้นตอน วิธีการที่จะจัดสร้างข้อสอบสามารถนำไปเผยแพร่กับเพื่อนร่วมงานได้”

ตารางที่ 14 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา ด้านผลลัพธ์ที่เกิดต่อองค์กร

ข้อที่	รายการประเมิน	$\bar{X}$	$SD$	ผลประเมิน
1	ประโยชน์ที่ได้รับสามารถบูรณาการสู่การทำงานเป็นทีมได้	4.15	0.64	มาก
2	ประโยชน์ที่ได้รับสามารถเพิ่มประสิทธิภาพการทำงาน	4.14	0.59	มาก
3	ประโยชน์ที่ได้รับทำให้ห้องค์กรยอมรับ	3.76	0.86	มาก
4	ความเป็นไปได้ในการปฏิรูประบบการทำงานในการปฏิบัติงาน	4.19	0.61	มาก
5	มีข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูงของนักเรียนในโรงเรียนสังกัดเมืองพัทยา	4.01	0.65	มาก
6	ผลการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-NET) ของนักเรียนในโรงเรียนสังกัดเมืองพัทยาสูงขึ้น	3.78	0.84	มาก

ตารางที่ 14 (ต่อ)

ข้อที่	รายการประเมิน	$\bar{X}$	$SD$	ผลประเมิน
7	ผลสัมฤทธิ์ทางการเรียนของนักเรียนในโรงเรียน สังกัดเมืองพัทยามีมาตรฐานที่ใกล้เคียงกัน	3.82	0.70	มาก
8	มีคลังข้อสอบสำหรับพัฒนาผลสัมฤทธิ์ทาง การเรียน	3.89	0.77	มาก
รวมเฉลี่ย		3.96	0.54	มาก

จากตารางที่ 14 พบว่า ผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ด้านผลลัพธ์ที่เกิดต่อองค์กร  
โดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 3.96, SD = 0.54$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า  
ความเป็นไปได้ในการปฏิรูประบบการทำงานในการปฏิบัติงาน มีคะแนนเฉลี่ยสูงที่สุด ( $\bar{X} = 4.19,$   
 $SD = 0.61$ ) รองลงมา ประโยชน์ที่ได้รับสามารถบูรณาการสู่การทำงานเป็นทีมได้ ( $\bar{X} = 4.15,$   
 $SD = 0.64$ ) ส่วนประโยชน์ที่ได้รับทำให้องค์กรยอมรับ มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.76,$   
 $SD = 0.86$ )

สรุปประเด็นความคิดเห็นในด้านประโยชน์ของโครงการสร้างข้อสอบมาตรฐานวัด  
ความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำนวน 17 คน 18  
ความคิดเห็น

ตารางที่ 15 สรุปความคิดเห็นในด้านประโยชน์ของโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา

สรุปจากความคิดเห็น	ตัวอย่างความคิดเห็น
1. ได้รับความรู้ เทคนิค และขั้นตอน การออกข้อสอบ	1. ได้เรียนรู้ เทคนิค วิธีการ การสร้างข้อสอบที่ได้ มาตรฐาน และเที่ยงตรง มีการวิเคราะห์ข้อสอบ เป็นรายข้อ 2. ได้รับความรู้เกี่ยวกับการออกข้อสอบและ การวิเคราะห์ข้อสอบ 3. ได้รับทราบเทคนิคในการทำข้อสอบ

## ตารางที่ 15 (ต่อ)

สรุปจากความคิดเห็น	ตัวอย่างความคิดเห็น
2. ได้นำความรู้จัดสร้างข้อสอบ	<ol style="list-style-type: none"> <li>1. ได้ลงมือปฏิบัติการสร้างข้อสอบให้ตรงตามวัตถุประสงค์การเรียนรู้ และตรงตามมาตรฐานตัวชี้วัด ทำให้มีความเข้าใจในการออกข้อสอบ</li> <li>2. ได้เรียนรู้วิธีการวิเคราะห์ตัวชี้วัด การออกข้อสอบตามตัวชี้วัดการทำ Test Blueprint การหาค่าความยากง่าย ซึ่งนำไปใช้ได้จริง</li> <li>3. ฝึกการทำข้อสอบหลากหลาย</li> </ol>
3. ได้นำความรู้มาพัฒนาจริง	<ol style="list-style-type: none"> <li>1. นำความรู้มาพัฒนางานในหน้าที่ให้มีประสิทธิภาพมากขึ้น</li> <li>2. ครูออกข้อสอบแนวใหม่แล้วสามารถนำมาใช้กับนักเรียนของตนเองได้</li> <li>3. ใช้ในการวัดความสามารถด้านการเรียนของผู้เรียนว่ามีความพร้อมในการทำข้อสอบมากน้อยเพียงใด และมีความเข้าใจในเนื้อหาการเรียนแค่ไหน</li> <li>4. ดีมาก และสามารถนำไปใช้ได้กับนักเรียน</li> </ol>

จากตารางที่ 15 พบว่า ความคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับประโยชน์ได้ของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่มโดย กลุ่มที่ 1 ด้านได้รับความรู้ เทคนิค และขั้นตอนการออกข้อสอบมีผู้ให้ความคิดเห็น ดังนี้ “คิดันคิดว่าได้รับความรู้เกี่ยวกับการออกข้อสอบและการวิเคราะห์ข้อสอบ, ได้เรียนรู้ เทคนิค วิธีการ การสร้างข้อสอบที่ได้มาตรฐาน และเที่ยงตรง มีการวิเคราะห์ข้อสอบเป็นรายข้อ” กลุ่มที่ 2 ด้านได้นำความรู้จัดสร้างข้อสอบมีผู้ให้ความคิดเห็น ดังนี้ “คิดันคิดว่าได้เรียนรู้วิธีการวิเคราะห์ตัวชี้วัด การออกข้อสอบตามตัวชี้วัดการทำ Test blueprint การหาค่าความยากง่าย ซึ่งนำไปใช้ได้จริง, ได้ลงมือปฏิบัติการสร้างข้อสอบให้ตรงตามวัตถุประสงค์การเรียนรู้ และตรงตามมาตรฐานตัวชี้วัดทำให้มีความเข้าใจในการออกข้อสอบและฝึกการทำข้อสอบหลากหลาย” กลุ่มที่ 3 ด้านได้นำความรู้มาพัฒนาจริงมีผู้ให้ความคิดเห็น ดังนี้ “คิดันคิดว่าครูออกข้อสอบแนวใหม่แล้วสามารถนำมาใช้กับนักเรียนของตนเองได้, ดีมาก และสามารถ

นำไปใช้ได้กับนักเรียน, ใช้ในการวัดความสามารถด้านการเรียนของผู้เรียนว่ามีความพร้อมในการทำข้อสอบมากน้อยเพียงใด และมีความเข้าใจในเนื้อหาการเรียนแค่ไหนและสามารถนำความรู้มาพัฒนางานในหน้าที่ให้มีประสิทธิภาพมากขึ้น

ตารางที่ 16 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานผลการประเมินโดยภาพรวมของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา รายนาม

ลำดับที่	รายการประเมิน	$\bar{X}$	$SD$	ผลประเมิน
1	การประเมินปฏิกิริยาตอบสนอง	3.70	0.47	มาก
2	การประเมินการเรียนรู้	4.05	0.50	มากที่สุด
3	การประเมินพฤติกรรม	4.01	0.63	มาก
4	การประเมินผลลัพธ์ที่เกิดต่อองค์กร	3.97	0.54	มาก
รวมเฉลี่ย		3.93	0.47	มาก

จากตารางที่ 16 พบว่า ผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 3.93, SD = 0.47$ ) เมื่อพิจารณารายด้านเรียงจากมากไปน้อย พบว่า ด้านการเรียนรู้มีคะแนนเฉลี่ยสูงที่สุด ( $\bar{X} = 4.05, SD = 0.50$ ) รองลงมา ด้านพฤติกรรม ( $\bar{X} = 4.01, SD = 0.63$ ) และด้านผลลัพธ์ที่เกิดต่อองค์กร ( $\bar{X} = 3.97, SD = 0.54$ ) ส่วนด้านการประเมินปฏิกิริยาตอบสนองมีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.70, SD = 0.47$ )

ตอนที่ 3 การวิเคราะห์เปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัด เมืองพัทยา  
โดยบูรณาการรูปแบบการประเมินของเคิร์กแพททริกและการประเมินอภิमान

ตารางที่ 17 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา  
จำแนกตามเพศ

ด้านที่	รายการ	ชาย		หญิง	
			<i>SD</i>	$\bar{X}$	<i>SD</i>
1	การประเมินปฏิกิริยาตอบสนอง	3.89	0.52	3.68	0.46
2	การประเมินการเรียนรู้	4.09	0.47	4.04	0.51
3	การประเมินพฤติกรรม	4.10	0.53	4.00	0.65
4	การประเมินผลลัพธ์ที่เกิดต่อองค์กร	4.13	0.46	3.95	0.54
	ค่าเฉลี่ยรวม	4.05	0.45	3.92	0.47

จากตารางที่ 17 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามเพศ  
นั้น โดยภาพรวมเพศชายมีความพึงพอใจมากกว่าเพศหญิง ( $\bar{X} = 4.05$ ) เมื่อแยกเป็นรายด้านปรากฏ  
ว่าทั้งด้านปฏิกิริยาตอบ ด้านการเรียนรู้ พฤติกรรม และผลลัพธ์ที่เกิดต่อองค์กร เพศชายมีความพึง  
พอใจมากกว่าเพศหญิง ( $\bar{X} = 3.89$ ), ( $\bar{X} = 4.09$ ), ( $\bar{X} = 4.10$ ) และ  $\bar{X} = 4.13$ ) ตามลำดับ



ตารางที่ 18 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยารายด้านและ  
โดยรวมของผู้เข้าอบรมที่มีเพศต่างกัน โดยการทดสอบค่าที (Independent *t*-test)

ตัวแปร	เพศ	<i>n</i>	$\bar{X}$	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig.</i>
ด้านปฏิบัติการ	ชาย	12	3.89	0.52	1.521	112	0.131
	หญิง	102	3.68	0.46			
ด้านทบทวน	ชาย	12	4.09	0.47	0.345	112	0.731
	หญิง	102	4.04	0.51			
ด้านพฤติกรรม	ชาย	12	4.10	0.53	0.514	112	0.608
	หญิง	102	4.00	0.65			
ด้านผลลัพธ์ที่เกิดต่อ องค์กร	ชาย	12	4.12	0.46	1.081	112	0.282
	หญิง	102	3.95	0.54			
รวม	ชาย	12	4.05	0.45	0.950	112	0.344
	หญิง	102	3.92	0.47			

จากตารางที่ 18 พบว่า ผู้เข้าอบรมในโครงการที่มีเพศต่างกัน มีความพึงพอใจต่อ  
โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียน  
ในสังกัดเมืองพัทยา ทั้งรายด้านทุกด้านและ โดยรวมต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ

ตารางที่ 19 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา  
จำแนกตามช่วงอายุ

ด้านที่	รายการ	20-30 ปี		31-40 ปี		41-50 ปี		51-60 ปี	
		$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$
1	การประเมินปฏิกริยา ตอบสนอง	3.60	0.55	3.67	0.55	3.53	0.57	3.77	0.48
2	การประเมินการเรียนรู้	4.13	0.41	4.06	0.50	4.01	0.60	4.03	0.43
3	การประเมินพฤติกรรม	4.10	0.58	4.11	0.59	3.93	0.55	4.14	0.54
4	การประเมินผลลัพธ์ที่เกิด ต่อองค์กร	4.00	0.35	4.05	0.61	3.84	0.58	4.16	0.52
	ค่าเฉลี่ยรวม	3.96	0.43	3.97	0.49	3.83	0.45	4.02	0.41

จากตารางที่ 19 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตามช่วงอายุนั้น โดยภาพรวมช่วงอายุระหว่าง 51-60 ปี มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.02$ ) รองลงมา ช่วงอายุระหว่าง 31-40 ปี ( $\bar{X} = 3.97$ ) และช่วงอายุระหว่าง 41-50 ปี ( $\bar{X} = 3.83$ ) มีความพึงพอใจน้อยสุด เมื่อแยกเป็นรายด้านสามารถแยกได้ ดังนี้ ด้านปฏิกริยาตอบสนองช่วงอายุระหว่าง 51-60 ปี มีความพึงพอใจมากที่สุด ( $\bar{X} = 3.77$ ) รองลงมา ช่วงอายุระหว่าง 31-40 ปี ( $\bar{X} = 3.67$ ) และช่วงอายุระหว่าง 41-50 ปี ( $\bar{X} = 3.53$ ) มีความพึงพอใจน้อยสุด ด้านการเรียนรู้ช่วงอายุระหว่าง 20-30 ปี มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.13$ ) รองลงมา ช่วงอายุระหว่าง 31-40 ปี ( $\bar{X} = 4.06$ ) และช่วงอายุระหว่าง 41-50 ปี ( $\bar{X} = 4.01$ ) มีความพึงพอใจน้อยสุด ด้านพฤติกรรมช่วงอายุระหว่าง 51-60 ปี มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.14$ ) รองลงมา ช่วงอายุระหว่าง 31-40 ปี ( $\bar{X} = 4.11$ ) และช่วงอายุระหว่าง 41-50 ปี ( $\bar{X} = 3.93$ ) มีความพึงพอใจน้อยสุด และด้านผลลัพธ์ที่เกิดต่อองค์กรช่วงอายุระหว่าง 51-60 ปี มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.16$ ) รองลงมา ช่วงอายุระหว่าง 31-40 ปี ( $\bar{X} = 4.05$ ) และช่วงอายุระหว่าง 41-50 ปี ( $\bar{X} = 3.84$ ) มีความพึงพอใจน้อยสุด

ตารางที่ 20 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา  
จำแนกตามประสบการณ์สอน

ด้านที่	รายการ	น้อยกว่า 10 ปี		11-20 ปี		21-30 ปี		31-40 ปี	
		$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$
1	การประเมินปฏิภิรียา ตอบสนอง	3.72	0.60	3.49	0.48	3.79	0.40	3.75	0.35
2	การประเมินการเรียนรูู้	4.15	0.53	3.89	0.47	3.99	0.42	4.44	0.62
3	การประเมิน พฤติกรรม	4.13	0.58	3.97	0.60	4.06	0.43	4.50	0.71
4	การประเมินผลลัพธ์ที่ เกิดต่อองค์กร	4.04	0.61	3.94	0.59	4.06	0.46	4.02	0.35
ค่าเฉลี่ยรวม		4.01	0.49	3.82	0.45	3.98	0.34	4.36	0.51

จากตารางที่ 20 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตาม  
ประสบการณ์ในการสอนนั้น โดยภาพรวมมีประสบการณ์การสอนระหว่าง 31-40 ปี  
มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.36$ ) รองลงมา ประสบการณ์การสอนน้อยกว่า 10 ปี ( $\bar{X} = 4.01$ )  
และประสบการณ์การสอนระหว่าง 11-20 ปี ( $\bar{X} = 3.82$ ) มีความพึงพอใจน้อยสุด เมื่อแยกเป็น  
รายด้านสามารถแยกได้ ดังนี้ ด้านปฏิภิรียาตอบสนองประสบการณ์การสอนระหว่าง 21-30 ปี  
มีความพึงพอใจมากที่สุด ( $\bar{X} = 3.79$ ) รองลงมา ประสบการณ์การสอนระหว่าง 31-40 ปี  
( $\bar{X} = 3.75$ ) และประสบการณ์การสอนระหว่าง 11-20 ปี ( $\bar{X} = 3.49$ ) มีความพึงพอใจน้อยสุด  
ด้านการเรียนรูู้ประสบการณ์การสอนระหว่าง 31-40 ปี มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.44$ )  
รองลงมา ประสบการณ์การสอนน้อยกว่า 10 ปี ( $\bar{X} = 4.15$ ) และประสบการณ์การสอนระหว่าง  
11-20 ปี ( $\bar{X} = 3.89$ ) มีความพึงพอใจน้อยสุด ด้านพฤติกรรมประสบการณ์การสอนระหว่าง  
31-40 ปี มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.50$ ) รองลงมา ประสบการณ์การสอนน้อยกว่า 10 ปี  
( $\bar{X} = 4.13$ ) และประสบการณ์การสอนระหว่าง 11-20 ปี ( $\bar{X} = 3.97$ ) มีความพึงพอใจน้อยสุด  
ด้านผลลัพธ์ที่เกิดต่อองค์กรประสบการณ์การสอนระหว่าง 21-30 ปี มีความพึงพอใจมากที่สุด

( $\bar{X} = 4.06$ ) รองลงมา ประสบการณ์การสอนน้อยกว่า 10 ปี ( $\bar{X} = 4.04$ ) และประสบการณ์การสอนระหว่าง 11-20 ปี ( $\bar{X} = 3.94$ ) มีความพึงพอใจน้อยสุด

ตารางที่ 21 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา จำแนกตามตำแหน่ง/ วิทยฐานะ

ด้านที่	รายการ	ครูผู้ช่วย		ครู		ครูชำนาญการ		ครูชำนาญการ/ ครูเชี่ยวชาญ		อื่น ๆ (เจ้าหน้าที่, ผู้บริหาร)	
		$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$
1	การประเมินปฏิบัติการ ตอบสนอง	3.20	0.27	3.65	0.48	3.63	0.52	3.83	0.26	3.76	0.67
2	การประเมิน การเรียนรู้	3.98	0.59	4.02	0.54	4.01	0.47	4.19	0.42	4.12	0.56
3	การประเมิน พฤติกรรม	3.90	0.22	4.10	0.53	4.03	0.56	4.25	0.42	4.11	0.68
4	การประเมินผลลัพธ์ ที่เกิดต่อองค์กร	3.80	0.45	4.00	0.59	3.96	0.55	4.42	0.49	4.09	0.67
	ค่าเฉลี่ยรวม	3.72	0.24	3.94	0.44	3.91	0.43	4.17	0.35	4.02	0.58

จากตารางที่ 21 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา จำแนกตามตำแหน่ง/ วิทยฐานะของผู้เข้าอบรมนั้น โดยภาพรวมครูชำนาญการ/ ครูเชี่ยวชาญ มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.17$ ) รองลงมา เจ้าหน้าที่และผู้บริหาร ( $\bar{X} = 4.02$ ) และครูผู้ช่วย ( $\bar{X} = 3.72$ ) มีความพึงพอใจน้อยสุด เมื่อแยกเป็นรายด้านสามารถแยกได้ ดังนี้ ด้านปฏิบัติการตอบสนองครูชำนาญการ/ ครูเชี่ยวชาญมีความพึงพอใจมากที่สุด ( $\bar{X} = 3.83$ ) รองลงมา เจ้าหน้าที่และผู้บริหาร ( $\bar{X} = 3.76$ ) และครูผู้ช่วย ( $\bar{X} = 3.20$ ) มีความพึงพอใจน้อยสุด ด้านการเรียนรู้ครูชำนาญการ/ ครูเชี่ยวชาญมีความพึงพอใจมากที่สุด ( $\bar{X} = 4.19$ ) รองลงมาเจ้าหน้าที่และผู้บริหาร ( $\bar{X} = 4.12$ ) และครูผู้ช่วย ( $\bar{X} = 3.98$ ) มีความพึงพอใจน้อยสุด ด้านพฤติกรรมครูชำนาญการ/ ครูเชี่ยวชาญ

มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.25$ ) รองลงมาเจ้าหน้าที่และผู้บริหาร ( $\bar{X} = 4.11$ ) และครูผู้ช่วย ( $\bar{X} = 3.90$ ) มีความพึงพอใจน้อยสุด ด้านผลลัพธ์ที่เกิดต่อองค์กรครูชำนาญการ/ ครูเชี่ยวชาญ มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.42$ ) รองลงมาเจ้าหน้าที่และผู้บริหาร ( $\bar{X} = 4.09$ ) และครูผู้ช่วย ( $\bar{X} = 3.82$ ) มีความพึงพอใจน้อยสุด

ตารางที่ 22 การเปรียบเทียบค่าเฉลี่ย ความพึงพอใจของผู้เข้าอบรมต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา รายด้านและโดยรวมของตำแหน่ง/ วิทยฐานะต่างกัน โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA)

แหล่งความแปรปรวน		<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>Sig.</i>
การประเมินปฏิภริยา ตอบสนอง	ระหว่างกลุ่ม	1.502	5	0.300	1.406	0.228
	ภายในกลุ่ม	23.068	108	0.214		
	รวม	24.570	113			
การประเมินการเรียนรู้	ระหว่างกลุ่ม	0.419	5	0.084	0.319	0.900
	ภายในกลุ่ม	28.339	108	0.262		
	รวม	28.758	113			
การประเมินพฤติกรรม	ระหว่างกลุ่ม	1.438	5	0.288	0.702	0.623
	ภายในกลุ่ม	44.229	108	0.410		
	รวม	45.667	113			
การประเมินผลลัพธ์ที่เกิด ต่อองค์กร	ระหว่างกลุ่ม	1.127	5	0.225	0.778	0.567
	ภายในกลุ่ม	31.181	108	0.290		
	รวม	32.408	113			
รวม	ระหว่างกลุ่ม	.868	5	0.174	0.779	0.567
	ภายในกลุ่ม	24.078	108	0.223		
	รวม	24.946	113			

จากตารางที่ 22 พบว่า ผู้เข้าอบรมในโครงการที่มีตำแหน่ง/ วิทยฐานะต่างกัน มีความพึงพอใจในโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง

ของนักเรียนโรงเรียนในสังกัดเมืองพัทยา ทั้งรายด้านทุกด้านและโดยรวมต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ

ตารางที่ 23 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา จำแนกตามระดับชั้นที่สอน

ด้านที่	รายการ	ประถมต้น		ประถมปลาย		มัธยมศึกษาตอนต้น		มัธยมศึกษาตอนปลาย		อื่นๆ (เจ้าหน้าที่, ผู้บริหาร)	
		$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$	$\bar{X}$	$SD$
1	การประเมินปฏิบัติการตอบสนอง	3.50	0.66	3.67	0.55	3.63	0.48	3.62	0.51	3.89	0.56
2	การประเมินการเรียนรู้	3.77	0.55	4.11	0.49	4.03	0.49	4.00	0.48	4.26	0.49
3	การประเมินพฤติกรรม	3.91	0.52	4.06	0.50	4.13	0.59	3.92	0.70	4.29	0.58
4	การประเมินผลลัพธ์ที่เกิดต่อองค์กร	3.84	0.57	4.04	0.54	4.01	0.64	3.85	0.59	4.32	0.50
	ค่าเฉลี่ยรวม	3.76	0.41	3.97	0.43	3.95	0.47	3.85	0.49	4.19	0.49

จากตารางที่ 23 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา จำแนกตามระดับชั้นที่สอนของผู้เข้าอบรมนั้น โดยภาพรวมเจ้าหน้าที่และผู้บริหาร มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.19$ ) รองลงมาประถมปลาย ( $\bar{X} = 3.97$ ) และประถมต้น ( $\bar{X} = 3.76$ ) มีความพึงพอใจน้อยสุด เมื่อแยกเป็นรายด้านสามารถแยกได้ ดังนี้ ด้านปฏิบัติการตอบสนองเจ้าหน้าที่และผู้บริหาร มีความพึงพอใจมากที่สุด ( $\bar{X} = 3.89$ ) รองลงมาประถมปลาย ( $\bar{X} = 3.67$ ) และประถมต้น ( $\bar{X} = 3.50$ ) มีความพึงพอใจน้อยสุด ด้านการเรียนรู้เจ้าหน้าที่และผู้บริหาร มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.26$ ) รองลงมาประถมปลาย ( $\bar{X} = 4.11$ ) และประถมต้น ( $\bar{X} = 3.77$ ) มีความพึงพอใจน้อยสุด ด้านพฤติกรรมเจ้าหน้าที่และผู้บริหาร มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.29$ ) รองลงมามัธยมศึกษาตอนต้น ( $\bar{X} = 4.13$ ) และประถมต้น ( $\bar{X} = 3.91$ ) มีความพึงพอใจน้อยสุด ด้านผลลัพธ์

ที่เกิดต่อองค์กรเจ้าหน้าที่และผู้บริหาร มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.32$ ) รองลงมาประณมปลาย ( $\bar{X} = 4.04$ ) และประณมต้น ( $\bar{X} = 3.84$ ) มีความพึงพอใจน้อยสุด

ตารางที่ 24 การเปรียบเทียบค่าเฉลี่ย ความพึงพอใจของผู้เข้าอบรมต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา รายด้านและโดยรวมของระดับชั้นที่สอนต่างกัน โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA)

แหล่งความแปรปรวน		<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>Sig.</i>
การประเมินปฏิภริยา ตอบสนอง	ระหว่างกลุ่ม	0.593	4	0.148	0.674	0.611
	ภายในกลุ่ม	23.977	109	0.220		
	รวม	24.570	113			
การประเมินการเรียนรู้	ระหว่างกลุ่ม	2.027	4	0.507	2.066	0.090
	ภายในกลุ่ม	26.731	109	0.245		
	รวม	28.756	113			
การประเมินพฤติกรรม	ระหว่างกลุ่ม	0.378	4	0.094	0.227	0.323
	ภายในกลุ่ม	45.290	109	0.416		
	รวม	45.667	113			
การประเมินผลลัพธ์ที่เกิด ต่อองค์กร	ระหว่างกลุ่ม	1.313	4	0.328	1.151	0.337
	ภายในกลุ่ม	31.094	109	.285		
	รวม	32.408	113			
รวม	ระหว่างกลุ่ม	0.734	4	0.184	0.827	0.511
	ภายในกลุ่ม	24.212	109	0.222		
	รวม	24.946	113			

จากตารางที่ 24 พบว่า ผู้เข้าร่วมการอบรมโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา ที่ระดับชั้นสอนต่างกัน มีค่าเฉลี่ยผลการประเมินของรายด้านและโดยรวมแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตารางที่ 25 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา  
จำแนกตามกลุ่มสาระการเรียนรู้

ด้านที่	รายการ	ไทย	คณิต	วิทย์	สังคม	สุขฯ	ศิลปะ	การงาน	ต่าง ประเทศ	อื่น ๆ
1	การประเมินปฏิภริยา ตอบสนอง	3.71	3.42	3.79	3.42	3.38	3.88	3.69	3.67	3.89
2	การประเมินการเรียนรู้	4.07	3.88	4.01	3.82	4.04	3.99	4.06	4.26	4.26
3	การประเมินพฤติกรรม	3.93	3.85	4.00	4.17	3.96	4.00	4.08	4.38	4.29
4	การประเมินผลลัพธ์ ที่เกิดต่อองค์กร	3.93	3.77	3.83	4.13	3.92	4.00	3.88	4.38	4.32
ค่าเฉลี่ยรวม		3.91	3.73	3.91	3.88	3.82	3.97	3.93	4.17	4.19

จากตารางที่ 25 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนก  
กลุ่มสาระการเรียนรู้ของผู้เข้าอบรมนั้น โดยภาพรวมเจ้าหน้าที่และผู้บริหาร มีความพึงพอใจ  
มากที่สุด ( $\bar{X} = 4.19$ ) รองลงมา กลุ่มสาระการเรียนรู้ภาษาต่างประเทศ ( $\bar{X} = 4.17$ ) และ  
กลุ่มสาระการเรียนรู้ ( $\bar{X} = 3.73$ ) มีความพึงพอใจน้อยสุด เมื่อแยกเป็นรายด้านสามารถแยกได้ ดังนี้  
ด้านปฏิภริยาตอบสนองเจ้าหน้าที่และผู้บริหาร มีความพึงพอใจมากที่สุด ( $\bar{X} = 89$ ) รองลงมา  
กลุ่มสาระการเรียนรู้ศิลปะ ( $\bar{X} = 3.88$ ) และกลุ่มสาระการเรียนรู้สุขศึกษาและพลศึกษา  
( $\bar{X} = 3.38$ ) มีความพึงพอใจน้อยสุด ด้านการเรียนรู้เจ้าหน้าที่และผู้บริหาร และกลุ่มสาระการเรียนรู้  
ภาษาต่างประเทศมีความพึงพอใจมากที่สุด ( $\bar{X} = 4.26$ ) รองลงมา กลุ่มสาระการเรียนรู้ภาษาไทย  
( $\bar{X} = 4.07$ ) และกลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม ( $\bar{X} = 3.88$ )  
มีความพึงพอใจน้อยสุด ด้านพฤติกรรมกลุ่มสาระการเรียนรู้ภาษาต่างประเทศ มีความพึงพอใจ  
มากที่สุด ( $\bar{X} = 4.38$ ) รองลงมา เจ้าหน้าที่และผู้บริหาร ( $\bar{X} = 4.29$ ) และกลุ่มสาระการเรียนรู้  
คณิตศาสตร์ ( $\bar{X} = 3.85$ ) มีความพึงพอใจน้อยสุด ด้านผลลัพธ์ที่เกิดต่อองค์กรกลุ่มสาระการเรียนรู้  
ภาษาต่างประเทศ มีความพึงพอใจมากที่สุด ( $\bar{X} = 4.38$ ) รองลงมา เจ้าหน้าที่และผู้บริหาร  
( $\bar{X} = 4.32$ ) และกลุ่มสาระการเรียนรู้คณิตศาสตร์ ( $\bar{X} = 3.77$ ) มีความพึงพอใจน้อยสุด



ตารางที่ 26 การเปรียบเทียบค่าเฉลี่ย ความพึงพอใจของผู้เข้าอบรมต่อโครงการสร้าง  
 ข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียน  
 ในสังกัดเมืองพัทยา รายด้านและโดยรวมของกลุ่มสาระการเรียนรู้ต่างกัน  
 โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA)

แหล่งความแปรปรวน		<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>Sig.</i>
การประเมินปฏิภริยา ตอบสนอง	ระหว่างกลุ่ม	1.125	8	0.141	0.630	0.751
	ภายในกลุ่ม	23.445	105	0.223		
	รวม	24.570	113			
การประเมินการเรียนรู้	ระหว่างกลุ่ม	2.228	8	0.279	1.102	0.368
	ภายในกลุ่ม	26.530	105	0.253		
	รวม	28.758	113			
การประเมินพฤติกรรม	ระหว่างกลุ่ม	2.847	8	0.356	0.873	0.542
	ภายในกลุ่ม	42.821	105	0.408		
	รวม	45.667	113			
การประเมินผลลัพธ์ที่เกิด ต่อองค์กร	ระหว่างกลุ่ม	3.484	8	0.436	1.581	0.139
	ภายในกลุ่ม	28.924	105	0.275		
	รวม	32.408	113			
รวม	ระหว่างกลุ่ม	1.769	8	0.221	1.002	0.439
	ภายในกลุ่ม	23.178	105	0.221		
	รวม	24.946	113			

จากตารางที่ 26 พบว่า ผู้เข้าร่วมการอบรมโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
 ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา ที่อยู่กลุ่มสาระการเรียนรู้  
 ต่างกันมีค่าเฉลี่ยผลการประเมินของรายด้านและโดยรวมแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ตารางที่ 27 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา  
จำแนกตามวุฒิการศึกษาสูงสุด

ด้านที่	รายการ	ปริญญาตรี		ปริญญาโท	
		$\bar{X}$	$SD$	$\bar{X}$	$SD$
1	การประเมินปฏิกิริยาตอบสนอง	3.71	0.38	3.68	0.57
2	การประเมินการเรียนรู้	4.07	0.53	4.02	0.47
3	การประเมินพฤติกรรม	4.05	0.58	3.95	0.71
4	การประเมินผลลัพธ์ที่เกิดต่อ องค์กร	3.93	0.54	4.02	0.53
	ค่าเฉลี่ยรวม	3.94	0.44	3.92	0.51

จากตารางที่ 27 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกตาม  
วุฒิการศึกษาสูงสุดนั้น โดยภาพรวมวุฒิปริญญาตรีมีความพึงพอใจมากกว่าวุฒิปริญญาโท  
( $\bar{X} = 3.95$ ) เมื่อแยกเป็นรายด้านสามารถแยกได้ ดังนี้ ด้านผลลัพธ์ที่เกิดต่อองค์กรวุฒิปริญญาโท  
มีความพึงพอใจมากกว่าวุฒิปริญญาตรี ( $\bar{X} = 4.02$ ) ส่วนด้านปฏิกิริยาตอบสนอง การเรียนรู้ และ  
พฤติกรรมวุฒิปริญญาตรีความพึงพอใจมากกว่าวุฒิปริญญาโท ( $\bar{X} = 3.71$ ), ( $\bar{X} = 4.07$ ) และ  
( $\bar{X} = 4.05$ ) ตามลำดับ

ตารางที่ 28 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา รายด้าน  
และโดยรวมของผู้เข้าอบรมที่มีวุฒิการศึกษาต่างกัน โดยการทดสอบค่าที  
(Independent *t*-test)

ตัวแปร	วุฒิทาง การศึกษาสูงสุด	<i>n</i>	$\bar{X}$	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig.</i>
ด้านปฏิกริยา	ปริญญาตรี	68	3.71	0.38	0.288	72.311	0.774
ตอบสนอง	ปริญญาโท	46	3.68	0.57			
ด้านการเรียนรู้	ปริญญาตรี	68	4.07	0.53	0.516	112	0.607
	ปริญญาโท	46	4.02	0.47			
ด้านพฤติกรรม	ปริญญาตรี	68	4.05	0.58	0.831	82.940	0.408
	ปริญญาโท	46	3.95	0.71			
ด้านผลลัพธ์ที่เกิด	ปริญญาตรี	68	3.93	0.54	-0.940	112	0.349
ต่อองค์กร	ปริญญาโท	46	4.02	0.53			
รวม	ปริญญาตรี	68	3.94	0.44	0.233	86.511	0.816
	ปริญญาโท	46	3.92	0.51			

\**P* < 0.05

จากตารางที่ 28 พบว่า ผู้เข้าอบรมในโครงการที่มีวุฒิการศึกษาต่างกัน มีความพึงพอใจ  
ต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน  
โรงเรียนในสังกัดเมืองพัทยา ทั้งรายด้านทุกด้านและ โดยรวมต่างกันอย่างไรไม่มีนัยสำคัญทางสถิติ

ตารางที่ 29 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา  
จำแนกวุฒิตามสาขา

ด้านที่	รายการ	สาขาการศึกษา		สาขาอื่น	
		$\bar{X}$	<i>SD</i>	$\bar{X}$	<i>SD</i>
1	การประเมินปฏิกิริยาตอบสนอง	3.61	0.38	4.51	0.40
2	การประเมินการเรียนรู้	3.99	0.49	4.55	0.40
3	การประเมินพฤติกรรม	3.94	0.62	4.67	0.30
4	การประเมินผลลัพธ์ที่เกิดต่อองค์กร	3.90	0.50	4.58	0.53
	ค่าเฉลี่ยรวม	3.86	0.43	4.58	0.24

จากตารางที่ 29 พบว่า การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐาน  
วัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา จำแนกวุฒิ  
ตามสาขาของผู้เข้าอบรมนั้น โดยภาพรวมผู้มีวุฒิสาखाอื่น ๆ มีความพึงพอใจต่อโครงการมากกว่า  
ผู้มีวุฒิทางการศึกษา ( $\bar{X} = 3.58$ ) เมื่อแยกเป็นรายด้านปรากฏว่าทั้งด้านปฏิกิริยาตอบสนอง  
ด้านการเรียนรู้ ด้านพฤติกรรม และด้านผลลัพธ์ที่เกิดต่อองค์กร วุฒิสาखाอื่น ๆ มีความพึงพอใจ  
มากกว่าวุฒิทางการศึกษา ( $\bar{X} = 4.15$ ) ( $\bar{X} = 4.55$ ), ( $\bar{X} = 4.67$ ) และ ( $\bar{X} = 4.58$ ) ตามลำดับ

ตารางที่ 30 การเปรียบเทียบความพึงพอใจต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้  
ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา รายด้าน  
และโดยรวมของผู้เข้าอบรมจำแนกวุฒิตามสาขา โดยการทดสอบค่าที  
(Independent *t*-test)

ตัวแปร	จำแนกวุฒิตามสาขา	<i>n</i>	$\bar{X}$	<i>SD</i>	<i>t</i>	<i>df</i>	<i>Sig</i>
ด้านปฏิกริยา	วุฒิทางการศึกษา	103	3.62	0.38	-7.345*	112	0.000
	วุฒิสาखाอื่น ๆ	11	4.51	0.40			
ด้านการเรียนรู้	วุฒิทางการศึกษา	103	3.99	0.49	-3.636*	112	0.000
	วุฒิสาखाอื่น ๆ	11	4.54	0.40			
ด้านพฤติกรรม	วุฒิทางการศึกษา	103	3.94	0.62	-6.704*	20.82	0.000
	วุฒิสาखाอื่น ๆ	11	4.67	0.30		3	
ด้านผลลัพธ์ที่ เกิดต่อองค์กร	วุฒิทางการศึกษา	103	3.90	0.50	-4.286*	112	0.000
	วุฒิสาखाอื่น ๆ	11	4.58	0.53			
รวม	วุฒิทางการศึกษา	103	3.86	0.43	-8.410*	17.75	0.000
	วุฒิสาखाอื่น ๆ	11	4.58	0.24		9	

\**p* < .05

จากตารางที่ 30 พบว่า ผู้เข้าอบรมในโครงการจำแนกวุฒิตามสาขา มีความพึงพอใจ  
ต่อโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียน  
ในสังกัดเมืองพัทยา เมื่อพิจารณารายด้านและโดยรวมมีความพึงพอใจต่อโครงการแตกต่างกัน  
อย่างมีนัยสำคัญทางสถิติที่ระดับ .05

#### ตอนที่ 4 การวิเคราะห์ความสัมพันธ์

ตารางที่ 31 ผลการวิเคราะห์ความสัมพันธ์ด้านปฏิริยาตอบสนอง (1), ด้านการเรียนรู้ (2), ด้านพฤติกรรม (3), ด้านผลลัพธ์ที่เกิดกับองค์กร (4) และด้านรวมกับอายุ และประสบการณ์ของผู้เข้าร่วมโครงการ โดยใช้สูตรสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

	อายุ	ประสบการณ์ สอน	ด้าน 1	ด้าน 2	ด้าน 3	ด้าน 4	ด้านรวม
อายุ	-						
ประสบการณ์สอน	.817**	-					
ด้าน 1	-.062	-.159	-				
ด้าน 2	-.026	-.130	.617**	-			
ด้าน 3	.022	-.040	.800**	.606**	-		
ด้าน 4	.020	-.053	.682**	.656**	.765**	-	
ด้านรวม	-.009	-.103	.879**	.814**	.917**	.889**	-

\*.Correlation is significant at the .05 level (2-tailed).

\*\*Correlation is significant at the .01 level (2-tailed).

จากตารางที่ 31 พบว่า อายุกับประสบการณ์สอน, ด้านปฏิริยากับด้านการเรียนรู้, ด้านปฏิริยาตอบสนองกับด้านพฤติกรรม, ด้านปฏิริยาตอบสนองกับด้านผลลัพธ์ที่เกิดต่อองค์กร, ด้านปฏิริยาตอบสนองกับด้านรวม, ด้านการเรียนรู้กับด้านพฤติกรรม, ด้านการเรียนรู้กับด้านผลลัพธ์ที่เกิดต่อองค์กร, ด้านการเรียนรู้กับด้านรวม, ด้านพฤติกรรมกับด้านผลลัพธ์ที่เกิดต่อองค์กร, ด้านพฤติกรรมกับด้านรวม, ด้านผลลัพธ์ที่เกิดต่อองค์กรกับด้านรวม มีความสัมพันธ์กันในทิศทางบวก อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

## บทที่ 5

### สรุป อภิปรายผลและข้อเสนอแนะ

การประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง ของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริก และการประเมินอภิมาน มีวัตถุประสงค์เพื่อประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบ การประเมินของเคิร์กแพตทริกและการประเมินอภิมาน โดยมีวัตถุประสงค์เฉพาะ เพื่อประเมิน ปฏิกริยาตอบสนอง (Reaction evaluation) โครงการตามรูปแบบการประเมินของเคิร์กแพตทริก และการประเมินอภิมาน เพื่อประเมินการเรียนรู้ (Learning evaluation) โครงการตามรูปแบบ การประเมินของเคิร์กแพตทริกและการประเมินอภิมาน เพื่อประเมินพฤติกรรม (Behavior evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน และเพื่อประเมิน ผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริกและ การประเมินอภิมาน โดยมีกลุ่มตัวอย่าง คือ บุคลากรที่เข้าอบรมในโครงการสร้างข้อสอบมาตรฐาน วัดความรู้ความสามารถในการคิดขั้นสูงของเมืองพัทยา จำนวน 114 คน ประกอบด้วย หัวหน้าฝ่าย วิชาการจำนวน 1 คน นักวิชาการศึกษา/ เจ้าหน้าที่โครงการจำนวน 3 คน ผู้อำนวยการสถานศึกษา/ รองผู้อำนวยการสถานศึกษา (รับผิดชอบงานวิชาการ) จำนวน 10 คน ผู้แทนครูวัดผลทุกสถานศึกษา จำนวน 11 คน ครูผู้สอนกลุ่มสาระการเรียนรู้ทั้ง 8 กลุ่มสาระการเรียนรู้จำนวน 89 คน เครื่องมือที่ใช้ ในการเก็บรวบรวมข้อมูล คือ แบบประเมิน โครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ ความสามารถในการคิดขั้นสูง สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ ค่าร้อยละ (Percentage) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (SD) สถิติการทดสอบค่าที (Independent t-test) สถิติ การวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA) การวิเคราะห์สัมประสิทธิ์ สหสัมพันธ์แบบเพียร์สัน (Pearson product moment correlation) และ การวิเคราะห์เนื้อหา (Content analysis)

### สรุปผลการวิจัย

การประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูง ของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและ การประเมินอภิมาน โดยภาพรวมอยู่ในระดับมาก ข้อมูลส่วนบุคคลของผู้ตอบแบบประเมินเมื่อ

จำแนกตามเพศของผู้เข้าอบรมโครงการส่วนใหญ่ เป็นเพศหญิง มีอายุอยู่ระหว่าง 31-40 ปี ประสบการณ์ที่สอนระหว่าง 1-10 ปี และมีตำแหน่งวิทยฐานะครูชำนาญการ สอนในระดับชั้นประถมศึกษา กลุ่มสาระการเรียนรู้ที่สอนภาษาไทย วุฒิการศึกษาสูงสุดปริญญาตรี และมีวุฒิทางการศึกษา ผลการประเมินความพึงพอใจโครงการอยู่ในระดับมาก ( $\bar{X} = 3.93, SD = 0.47$ ) เมื่อพิจารณารายด้านเรียงจากมากไปน้อย พบว่า ด้านการเรียนรู้มีคะแนนเฉลี่ยสูงที่สุด ( $\bar{X} = 4.05, SD = 0.50$ ) รองลงมา ด้านพฤติกรรม ( $\bar{X} = 4.01, SD = 0.63$ ) และด้านผลลัพธ์ที่เกิดต่อองค์กร ( $\bar{X} = 3.97, SD = 0.54$ ) ส่วนด้านการประเมินปฏิริยาตอบสนองมีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.70, SD = 0.47$ ) ซึ่งสรุปตามวัตถุประสงค์เฉพาะ ดังนี้

1. ผลการประเมินปฏิริยาตอบสนอง (Reaction evaluation) โครงการตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน

ผลการประเมินโดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 3.70, SD = 0.47$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า ความเหมาะสมเพียงพอของสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในการฝึกอบรมมีคะแนนเฉลี่ยสูงที่สุด ( $\bar{X} = 4.08, SD = 0.63$ ) รองลงมา ได้แก่ ความเหมาะสมของความพร้อมอุปกรณ์โสตทัศนูปกรณ์ ( $\bar{X} = 3.99, SD = 0.52$ ) ส่วนความเหมาะสมของวัน-เวลาในการจัดโครงการ มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.30, SD = 0.91$ )

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความเหมาะสมของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่มโดย กลุ่มที่ 1 ด้านระยะเวลาในการอบรมมีผู้ให้ความคิดเห็นทั้งในด้านบวกและด้านลบ ด้านบวก ดังนี้ “คิดว่าระยะเวลาในการอบรมครั้งนี้เหมาะสมดี” ในด้านลบ “คิดว่าการอบรมมีจำนวนระยะหลายช่วงเวลาทำให้กระทบกับการสอนของครู” กลุ่มที่ 2 ด้านสถานที่ในการจัดการอบรมมีผู้ให้ความคิดเห็นทั้งด้านบวกและด้านลบ ในด้านบวก ดังนี้ “คิดว่าสถานที่ในการอบรมดี สะดวกต่อผู้เข้าร่วมการอบรมเพราะติดถนนสุขุมวิท การเดินทางสะดวก” ในด้านลบ “คิดว่าบางครั้งเครียดเกินไปและจัดตรงกับวันที่มีนักเรียนมาเรียนจึงทำให้บางครั้งมีเสียงมารบกวนเวลาทำกิจกรรม” กลุ่มสุดท้ายด้านสิ่งอำนวยความสะดวกมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าดี มีสิ่งอำนวยความสะดวกเพียงพอต่อความต้องการและเหมาะสม”

2. ผลการประเมินการเรียนรู้ (Learning evaluation) โครงการตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน

ผลการประเมินโดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 4.05, SD = 0.50$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า ความรู้ ความเข้าใจในเรื่องนี้หลังการอบรมมีคะแนนเฉลี่ยสูงที่สุด ( $\bar{X} = 4.22, SD = 0.56$ ) รองลงมา ได้แก่ สามารถประมวลความคิดสู่การพัฒนางานอย่างเป็นระบบ



กับข้อ 6 สามารถปฏิบัติการออกแบบทดสอบวัดความรู้ความสามารถขั้นสูงของนักเรียนได้ ( $\bar{X} = 4.11, SD = 0.59$ ) ส่วนความเกิดทักษะและประสบการณ์ใหม่ ๆ จากการฝึกอบรม มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.91, SD = 0.65$ )

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความถูกต้องด้านเนื้อหาของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านเนื้อหาให้ผู้ให้ความคิดเห็นทั้งในด้านบวกและด้านลบ ด้านบวก ดังนี้ “คิดว่าความถูกต้องด้านเนื้อหาอยู่ในระดับดีมาก มีการนำเสนอกระแสมากปี พ.ศ. 2551 มาออกข้อสอบในแต่ละตัวชี้วัด” ด้านลบ ดังนี้ “คิดว่ามีเนื้อหาเยอะเข้าใจยาก” กลุ่มที่ 2 ด้านความรู้จากวิทยากรมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าวิทยากรอธิบายเข้าใจง่ายและวิทยากรประจำกลุ่มให้การช่วยเหลือดี” และกลุ่มที่ 3 ด้านการนำเสนอข้อมูลมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าควรมีการจัดลำดับการนำเสนอข้อมูลต่อผู้เข้ารับการอบรม”

3. ผลการประเมินพฤติกรรม (Behavior evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิमान

ผลการประเมินโดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 4.01, SD = 0.64$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า สามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการปฏิบัติงานได้มีคะแนนเฉลี่ยสูงที่สุด ( $\bar{X} = 4.18, SD = 0.65$ ) รองลงมา ได้แก่ สามารถให้คำปรึกษาแก่เพื่อนร่วมงานได้ ( $\bar{X} = 4.06, SD = 0.72$ ) ส่วนสามารถนำความรู้ไปเผยแพร่/ ถ่ายทอดแก่เพื่อนร่วมงานได้ มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.91, SD = 0.84$ )

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความเป็นไปได้หลังจากเข้าอบรม โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านมีข้อสอบที่มาตรฐานมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าจากการที่ได้ปฏิบัติกิจกรรมทำให้ได้ข้อสอบที่มีคุณภาพมากขึ้นในแต่ละรายวิชาที่สามารถจะวัดผู้เรียนได้ตรงตามมาตรฐานตัวชี้วัด” กลุ่มที่ 2 ด้านการพัฒนาตนเองมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าเข้าใจในการทำข้อสอบเพิ่มมากขึ้นและได้มีเกณฑ์มาตรฐานในการทำข้อสอบมากกว่าที่ควรพร้อมทั้งยังสามารถนำไปปรับปรุงข้อสอบของตนเองได้” กลุ่มที่ 3 ด้านการเผยแพร่ความรู้แก่เพื่อนร่วมงานมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าสามารถให้คำปรึกษาแก่เพื่อนร่วมงานในด้านการวิเคราะห์ตัวชี้วัดเพื่อจัดทำแผนผังข้อสอบ (Test blueprint) และการวิเคราะห์ข้อสอบรายข้อได้”

4. ผลการประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิमान

ผลการประเมินโดยภาพรวมอยู่ในระดับมาก ( $\bar{X} = 3.96, SD = 0.54$ ) เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า ความเป็นไปได้ในการปฏิรูประบบการทำงานในการปฏิบัติงาน มีคะแนนเฉลี่ยสูงสุดที่สุด ( $\bar{X} = 4.19, SD = 0.61$ ) รองลงมา ประโยชน์ที่ได้รับสามารถบูรณาการสู่การทำงานเป็นทีมได้ ( $\bar{X} = 4.15, SD = 0.64$ ) ส่วนประโยชน์ที่ได้รับทำให้องค์กรยอมรับ มีคะแนนเฉลี่ยต่ำสุด ( $\bar{X} = 3.76, SD = 0.86$ )

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับประโยชน์ของโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยาสามารถจำแนกได้เป็น 3 กลุ่มโดย กลุ่มที่ 1 ด้านได้รับความรู้ เทคนิค และขั้นตอนการออกข้อสอบมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าได้รับความรู้เกี่ยวกับเทคนิค วิธีการ การสร้างข้อสอบที่ได้มาตรฐาน และเที่ยงตรง มีการวิเคราะห์ข้อสอบเป็นรายข้อ” กลุ่มที่ 2 ด้านได้นำความรู้จัดสร้างข้อสอบมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าได้เรียนรู้วิธีการวิเคราะห์ตัวชี้วัด การออกข้อสอบตามตัวชี้วัดการทำ Test blueprint การหาค่าความยากง่าย ซึ่งนำไปใช้ได้จริง” กลุ่มที่ 3 ด้านได้นำความรู้มาพัฒนาจริงมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าครูออกข้อสอบแนวใหม่แล้วสามารถนำมาใช้กับนักเรียนของตนเองได้”

## อภิปรายผล

จากผลการประเมิน โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของ เคิร์กแพตทริกและการประเมินอภิमान สามารถอภิปรายผลได้ ดังนี้

1. ผลการประเมินปฏิกิริยาตอบสนอง (Reaction evaluation) โครงการตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิमान พบว่า โดยภาพรวมอยู่ในระดับมาก เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า ความเหมาะสมเพียงพอของสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในการฝึกอบรมมีคะแนนเฉลี่ยสูงสุด รองลงมา ได้แก่ ความเหมาะสมของความพร้อมอุปกรณ์โสตทัศนูปกรณ์ ส่วนความเหมาะสมของวัน-เวลาในการจัดโครงการ มีคะแนนเฉลี่ยต่ำสุด ทั้งนี้อาจเป็นเพราะว่าระยะเวลาการอบรมโครงการใช้เวลานาน สอดคล้องกับงานวิจัยของ จุฑารัตน์ หนันไชย (2552) ได้ศึกษาเรื่อง การประยุกต์ใช้แบบจำลองของเคิร์กแพตทริกในการประเมินและติดตามผล โครงการฝึกอบรมการแก้ไขปัญหายาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด พบว่า ชั้นปฏิกิริยาผู้เข้าร่วมโครงการมีความพึงพอใจในการดำเนินการฝึกอบรมอยู่ในระดับมาก ในส่วนของข้อเด่นในการใช้แบบจำลองของเคิร์กแพตทริก คือ สามารถนำมาประยุกต์ใช้ได้กับเด็กและเยาวชนกลุ่มนี้ได้อย่างเหมาะสม ในส่วนของข้อจำกัดมี 2 ลักษณะ คือ ข้อจำกัดที่เกิดจาก

เครื่องมือ และจากกระบวนการประยุกต์ใช้เครื่องมือในเรื่องสภาพแวดล้อมที่ไม่เอื้ออำนวย และอาจเป็นเพราะโครงการนี้

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความเหมาะสมของโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน โรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านระยะเวลาในการอบรมมีผู้ให้ความคิดเห็นทั้งในด้านบวกและด้านลบ ด้านบวก ดังนี้ “คิดว่าระยะเวลาในการอบรมครั้งนี้เหมาะสมดี” ในด้านลบ “คิดว่าการอบรมมีจำนวนระยะเวลาทำให้กระทบกับการสอนของครู” กลุ่มที่ 2 ด้านสถานที่ในการจัดการอบรมมีผู้ให้ความคิดเห็นทั้งด้านบวกและด้านลบ ในด้านบวก ดังนี้ “คิดว่าสถานที่ในการอบรมดี สะดวกต่อผู้เข้าร่วมการอบรมเพราะติดถนนสุขุมวิท การเดินทางสะดวก” ในด้านลบ “คิดว่าบางครั้งเครียดเกินไปและจัดตรงกับวันที่มีนักเรียนมาเรียนจึงทำให้บางครั้งมีเสียงมารบกวนเวลาทำกิจกรรม” กลุ่มสุดท้ายด้านสิ่งอำนวยความสะดวกมีผู้ให้ความคิดเห็น ดังนี้ “คิดว่าดี มีสิ่งอำนวยความสะดวกเพียงพอต่อความต้องการและเหมาะสม”

2. ผลการประเมินการเรียนรู้ (Learning evaluation) โครงการตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน พบว่า โดยภาพรวมอยู่ในระดับมาก เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า ความรู้ ความเข้าใจในเรื่องนี้หลังการอบรมมีคะแนนเฉลี่ยสูงที่สุด รองลงมา ได้แก่ สามารถประมวลความคิดสู่การพัฒนาอย่างเป็นระบบกับข้อ 6 สามารถปฏิบัติการออกแบบทดสอบวัดความรู้ความสามารถขั้นสูงของนักเรียนได้ ส่วนความเกิดทักษะและประสบการณ์ใหม่ ๆ จากการฝึกอบรม มีคะแนนเฉลี่ยต่ำสุด สอดคล้องกับงานวิจัยของ ศันสนีย์ อุดมอ่าง (2554) ได้ประเมินโครงการ การฝึกอบรมเชิงปฏิบัติการการผลิตทองม้วนเพื่อกระตุ้นเศรษฐกิจชุมชนตามแนวเศรษฐกิจพอเพียง พบว่า ผู้เข้าอบรมมีความพึงพอใจต่อการอบรมระดับมากถึงมากที่สุด ผู้เข้าอบรมมีความรู้ความเข้าใจ ได้คะแนนหลังการอบรมเฉลี่ยสูงกว่าคะแนนก่อนการอบรมเฉลี่ย **ชญญาภัค วงศ์บา (2555)** ได้ศึกษาเรื่องประเมินผลการฝึกอบรมตามโครงการ Champion โดยประยุกต์ใช้แนวทางการประเมินของเคิร์กแพตทริก ผลการวิจัย พบว่า ผู้เข้ารับการฝึกอบรมส่วนใหญ่สังกัดหน่วยงานราชการ เป็นเพศหญิง ระดับการศึกษาสูงสุด จบปริญญาตรี อายุเฉลี่ย 41.53 ส่วนเบี่ยงเบนมาตรฐาน 8.31 การประเมินระดับที่ 2 คะแนนเฉลี่ยก่อนการฝึกอบรม 15.71 ค่าส่วนเบี่ยงเบนมาตรฐาน 2.57 และคะแนนเฉลี่ยหลังการฝึกอบรม 25.02 ค่าส่วนเบี่ยงเบนมาตรฐาน 1.79 และยังสอดคล้องกับงานวิจัยของ จุฑารัตน์ ทน้นไชย (2552) ได้ศึกษาเรื่อง การประยุกต์ใช้แบบจำลองของเคิร์กแพตทริกในการประเมินและติดตามผลโครงการฝึกอบรมการแก้ไขปัญหาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด มีวัตถุประสงค์เพื่อศึกษาแบบจำลองของเคิร์กแพตทริกในการประเมินและติดตามผลโครงการ

ฝึกรอบรม ข้อเด่นและข้อจำกัดของการใช้แบบในการประยุกต์ใช้ในการประเมินสัมฤทธิ์ผลของโครงการแก้ไขปัญหายาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด พบว่า ในชั้นการเรียนรู้ผู้เข้าร่วมโครงการมีความรู้เกี่ยวกับโรคเอดส์และยาเสพติดในระดับสูง ในส่วนของข้อเด่นในการใช้แบบจำลองของเคิร์กแพตทริก คือ สามารถนำมาประยุกต์ใช้ได้กับเด็กและเยาวชนกลุ่มนี้อย่างเหมาะสม ในส่วนของข้อจำกัดมี 2 ลักษณะ คือ ข้อจำกัดที่เกิดจากเครื่องมือ และจากกระบวนการประยุกต์ใช้เครื่องมือในเรื่องสภาพแวดล้อมที่ไม่เอื้ออำนวย

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความถูกต้องด้านเนื้อหาของโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านเนื้อหาให้ผู้ให้ความคิดเห็นทั้งในด้านบวกและด้านลบ ด้านบวก ดังนี้ “ฉันคิดว่าความถูกต้องด้านเนื้อหาอยู่ในระดับดีมาก การให้ความรู้ในรายละเอียดของการออกข้อสอบจากคณะวิทยากรให้คำชี้แนะ แก้ไขและปรับปรุง, มีการนำสาระแกนกลางปี 2551 มาออกข้อสอบในแต่ละตัวชี้วัด” ด้านลบ ดังนี้ “ฉันคิดว่าเป็นโครงการที่ดี แต่มีเนื้อหาเยอะเข้าใจยาก, เนื้อหาในการบรรยายเยอะ” กลุ่มที่ 2 ด้านความรู้จากวิทยากรมีผู้ให้ความคิดเห็น ดังนี้ “ฉันคิดว่าวิทยากรอธิบายเข้าใจง่ายและวิทยากรประจำกลุ่มให้การช่วยเหลือดี” และกลุ่มที่ 3 ด้านการนำเสนอข้อมูลมีผู้ให้ความคิดเห็น ดังนี้ “ฉันคิดว่าควรมีการจัดลำดับการนำเสนอข้อมูลต่อผู้เข้ารับการอบรม, มีการนำเสนอลำช้าผู้ปฏิบัติเกิดการสับสนและต้องแก้ไขข้อมูลใหม่ทำให้การทำงานล่าช้า

3. ผลการประเมินพฤติกรรม (Behavior evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิमान พบว่า โดยภาพรวมอยู่ในระดับมาก เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า สามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการปฏิบัติงานได้มีคะแนนเฉลี่ยสูงที่สุด รองลงมา ได้แก่ สามารถให้คำปรึกษาแก่เพื่อนร่วมงานได้ ส่วนสามารถนำความรู้ไปเผยแพร่/ถ่ายทอดแก่เพื่อนร่วมงานได้ มีคะแนนเฉลี่ยต่ำสุด สอดคล้องกับงานวิจัยของ จุฑารัตน์ ทนัฒไชย (2552) ได้ศึกษาเรื่อง การประยุกต์ใช้แบบจำลองของเคิร์กแพตทริกในการประเมินและติดตามผลโครงการฝึกรอบรมการแก้ไขปัญหายาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด มีวัตถุประสงค์เพื่อศึกษาแบบจำลองของเคิร์กแพตทริกในการประเมินและติดตามผลโครงการฝึกรอบรม ข้อเด่นและข้อจำกัดของการใช้แบบในการประยุกต์ใช้ในการประเมินสัมฤทธิ์ผลของโครงการแก้ไขปัญหายาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด พบว่า ชั้นพฤติกรรมและชั้นผลลัพธ์ พบว่า ผู้เข้าร่วมโครงการมีการเปลี่ยนแปลงพฤติกรรมภายหลังจากการฝึกรอบรมในบางพฤติกรรม และสามารถนำผลที่ได้รับจากการฝึกรอบรมเรื่องยาเสพติดไปใช้ประโยชน์มากที่สุด ในส่วนของข้อเด่นในการใช้แบบจำลองของเคิร์กแพตทริก คือ สามารถนำมา

ประยุกต์ใช้ได้กับเด็กและเยาวชนกลุ่มนี้เหมาะสม ในส่วนของข้อจำกัดมี 2 ลักษณะ คือ ข้อจำกัดที่เกิดจากเครื่องมือ และจากกระบวนการประยุกต์ใช้เครื่องมือในเรื่องสภาพแวดล้อมที่ไม่เอื้ออำนวย

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับความเป็นไปได้ สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านมีข้อสอบที่มาตรฐานมีผู้ให้ความคิดเห็น ดังนี้ “ฉันคิดว่าจากการที่ได้ปฏิบัติกิจกรรมทำให้ได้ข้อสอบที่มีคุณภาพมากขึ้นในแต่ละรายวิชาที่สามารถจะวัดผู้เรียนได้ตรงตามมาตรฐานตัวชี้วัด, มีการสร้างข้อสอบที่เป็นระบบและได้มาตรฐานตรงตามตัวชี้วัด, สามารถนำข้อสอบที่สร้างไปใช้ทดสอบกับนักเรียนได้จริง” กลุ่มที่ 2 ด้านการพัฒนาตนเองมีผู้ให้ความคิดเห็น ดังนี้ “ฉันคิดว่าเข้าใจในการทำข้อสอบเพิ่มมากขึ้นและได้มีเกณฑ์มาตรฐานในการทำข้อสอบมากเท่าที่ควรพร้อมทั้งยังสามารถนำไปปรับปรุงข้อสอบของตนเองได้, มีการพัฒนาการจัดกิจกรรมการเรียนรู้ การสอนของครูให้สอดคล้องตรงตามแผนจัดการเรียนรู้ของมาตรฐานและตัวชี้วัดและนำความรู้ที่ได้รับมาปรับปรุงข้อสอบที่ใช้วัดกับเด็กนักเรียนและพัฒนาการเรียนการสอนได้” กลุ่มที่ 3 ด้านการเผยแพร่ความรู้แก่เพื่อนร่วมงานมีผู้ให้ความคิดเห็น ดังนี้ “ฉันคิดว่าสามารถให้คำปรึกษาแก่เพื่อนร่วมงานในด้านการวิเคราะห์ตัวชี้วัดเพื่อจัดทำแผนผังข้อสอบ (Test blueprint) และการวิเคราะห์ข้อสอบรายข้อได้, รู้ขั้นตอน วิธีการที่จะจัดสร้างข้อสอบสามารถนำไปเผยแพร่กับเพื่อนร่วมงานได้”

4. ผลการประเมินผลลัพธ์ที่เกิดต่อองค์กร (Result evaluation) ตามรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอภิมาน พบว่า โดยภาพรวมอยู่ในระดับมาก เมื่อพิจารณารายข้อเรียงจากมากไปน้อย พบว่า ความเป็นไปได้ในการปฏิรูประบบการทำงานในการปฏิบัติงาน มีคะแนนเฉลี่ยสูงสุด รองลงมาประโยชน์ที่ได้รับสามารถบูรณาการสู่การทำงานเป็นทีมได้ ส่วนประโยชน์ที่ได้รับทำให้องค์กรยอมรับ มีคะแนนเฉลี่ยต่ำสุด สอดคล้องกับงานวิจัยของ จุฑารัตน์ ทนัฒไชย (2552) ได้ศึกษาเรื่อง การประยุกต์ใช้แบบจำลองของเคิร์กแพตทริก ในการประเมินและติดตามผลโครงการฝึกอบรมการแก้ไขปัญหาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด มีวัตถุประสงค์เพื่อศึกษาแบบจำลองของเคิร์กแพตทริก ในการประเมินและติดตามผลโครงการฝึกอบรม ข้อเด่นและข้อจำกัดของการใช้แบบในการประยุกต์ใช้ในการประเมินสัมฤทธิ์ผลของโครงการแก้ไขปัญหาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด พบว่า ใน ชั้นพฤติกรรมและชั้นผลลัพธ์ พบว่า ผู้เข้าร่วมโครงการ มีการเปลี่ยนแปลงพฤติกรรมภายหลังจากการฝึกอบรมในบางพฤติกรรม และสามารถนำผลที่ได้รับจากการฝึกอบรมเรื่องยาเสพติดไปใช้ประโยชน์มากที่สุด ในส่วนของข้อเด่นในการใช้แบบจำลองของเคิร์กแพตทริก คือ สามารถนำมาประยุกต์ใช้ได้กับเด็กและเยาวชนกลุ่มนี้เหมาะสม

ในส่วนของคุณ้อจำกัดมี 2 ลักษณะ คือ ข้อจำกัดที่เกิดจากเครื่องมือ และจากกระบวนการประยุกต์ใช้เครื่องมือในเรื่องสภาพแวดล้อมที่ไม่เอื้ออำนวย

สรุปข้อคิดเห็นของกลุ่มตัวอย่างเกี่ยวกับประโยชน์ สามารถจำแนกได้เป็น 3 กลุ่ม โดย กลุ่มที่ 1 ด้านได้รับความรู้ เทคนิค และขั้นตอนการออกข้อสอบมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าได้รับความรู้เกี่ยวกับการออกข้อสอบและการวิเคราะห์ข้อสอบ, ได้เรียนรู้ เทคนิค วิธีการ การสร้างข้อสอบที่ได้มาตรฐาน และเที่ยงตรง มีการวิเคราะห์ข้อสอบเป็นรายข้อ” กลุ่มที่ 2 ด้านได้นำความรู้จัดสร้างข้อสอบมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าได้เรียนรู้วิธีการวิเคราะห์ ตัวชี้วัด การออกข้อสอบตามตัวชี้วัดการทำ Test blueprint การหาค่าความยากง่าย ซึ่งนำไปใช้ได้จริง, ได้ลงมือปฏิบัติการสร้างข้อสอบให้ตรงตามวัตถุประสงค์การเรียนรู้ และตรงตามมาตรฐาน ตัวชี้วัดทำให้มีความเข้าใจในการออกข้อสอบและฝึกการทำข้อสอบหลากหลาย” กลุ่มที่ 3 ด้านได้นำความรู้มาพัฒนาจริงมีผู้ให้ความคิดเห็น ดังนี้ “ดิฉันคิดว่าครูออกข้อสอบแนวใหม่แล้วสามารถนำมาใช้กับนักเรียนของตนเองได้, ดีมาก และสามารถนำไปใช้ได้กับนักเรียน, ใช้ในการวัดความสามารถด้านการเรียนของผู้เรียนว่ามีความพร้อมในการทำข้อสอบมากน้อยเพียงใด และมีความเข้าใจในเนื้อหาการเรียนแค่ไหนและสามารถนำความรู้มาพัฒนางานในหน้าที่ให้มีประสิทธิภาพมากขึ้น

### ข้อเสนอแนะ

โครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา เป็นโครงการที่พัฒนาความสามารถของพนักงานครูสังกัดเมืองพัทยา ในการพัฒนาข้อสอบที่มีมาตรฐาน จากการท้าวิจัย พบว่า จุดแข็งของโครงการนี้ คือ ได้ข้อสอบที่มีคุณภาพและมีมาตรฐาน บุคลากรได้รับความรู้ เทคนิควิธีการต่าง ๆ ในการออกข้อสอบ และมีคลังข้อสอบสำหรับพัฒนาผลสัมฤทธิ์ทางการเรียน และจุดอ่อนของโครงการ คือ ด้านระยะเวลาในการจัดการอบรม ความรู้ก่อนเข้าอบรมของผู้อบรมซึ่งไม่ได้มีการทดสอบก่อน และวิทยากรที่ให้การอบรมรวมทั้งวิทยากรพี่เลี้ยงไม่ทั่วถึง ซึ่งจากการท้าวิจัยมีจุดที่ควรพัฒนา ดังนี้

1. ควรให้มีการปรับวัน ระยะเวลาให้กระชับและเหมาะสม เนื่องจากโครงการนี้ใช้เวลาในการอบรมหลายระยะ และในแต่ละระยะใช้เวลาหลายวันทำให้กระทบต่อการจัดการเรียนการสอนของผู้อบรมและทำให้ผู้เข้าอบรมไม่ได้หยุดพัก
2. ควรให้ความรู้หลักการสร้างข้อสอบที่ถูกต้องทั้งแบบปรนัยและอัตนัยแก่ผู้เข้ารับการอบรมก่อนอบรมและควรเลือกผู้เข้าอบรมตรงกับชั้นที่ออกข้อสอบ เช่น โครงการออกข้อสอบในระดับชั้นประถมศึกษาปีที่ 2, ชั้นประถมศึกษาปีที่ 5 และชั้นมัธยมศึกษาปีที่ 2 ควรเลือก

ผู้เข้าอบรมที่สอนในระดับชั้นนั้น ๆ เพื่อให้เข้าใจในกระบวนการ เนื้อหาสาระ วิธีการ ในการออกข้อสอบ เนื่องจากครูบางท่านสอนหลายวิชา บางท่านไม่ได้สอนในชั้นที่ออกข้อสอบ ในระยะเวลานั้น

3. ควรเพิ่มวิทยากร โดยเฉพาะวิทยากรที่เลี้ยงประจำกลุ่มสาระการเรียนรู้ เพื่อจะได้ ให้ข้อเสนอแนะและดูแลได้อย่างทั่วถึง

#### ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ควรมีการประเมินถึงผลลัพธ์ (Outcomes) หลังจากเสร็จโครงการว่า มีข้อดี ข้อเสีย ของโครงการส่งผลต่อผู้เกี่ยวข้องอย่างไรในระยะยาว และผลกระทบ (Impact) ที่จะเกิดขึ้น ตามมาอย่างต่อเนื่อง ตลอดจนแนวทางการวางแผนโครงการในวงรอบต่อไปที่ผู้บริหารจะใช้เป็น สารสนเทศประกอบในการตัดสินใจว่าจะดำเนินการโครงการอย่างไร

2. ควรมีการวิจัยพัฒนาธนาคารข้อสอบ (Item bank) เพื่อให้ได้ข้อสอบที่มี ประสิทธิภาพและมีมาตรฐานมาใช้กับนักเรียนเพื่อพัฒนาผลสัมฤทธิ์ทางการเรียน และพัฒนา ผลการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-NET) ต่อไป

3. การประเมินโครงการ โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและ การประเมินอภิมาน อาจสามารถนำไปดัดแปลงหรือประยุกต์ใช้ในการประเมินโครงการอื่น ๆ

## บรรณานุกรม

- กำจัด สุดโต. (2553). การประเมินโครงการ โรงเรียนส่งเสริมสุขภาพของโรงเรียนสังกัดเทศบาลจังหวัดสุพรรณบุรี. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต, สาขาวิชาวิจัยและประเมินผลการศึกษา, คณะครุศาสตร์, มหาวิทยาลัยราชภัฏกาญจนบุรี.
- ขวัญภา อุนหกันต์. (2553). การประเมินโครงการคุณธรรมนำความรู้สู่ความดีของโรงเรียนต้นแบบคุณธรรมนำความรู้สู่ความดีสังกัดสำนักงานเขตพื้นที่การศึกษาสุพรรณบุรีเขต 2. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต, สาขาวิชาวิจัยและประเมินผลการศึกษา, คณะครุศาสตร์, มหาวิทยาลัยราชภัฏกาญจนบุรี.
- จุฑารัตน์ ทนน์ไชย. (2552). การประยุกต์ใช้แบบจำลองของเคิร์กแพทริกในการประเมินและติดตามผลโครงการฝึกอบรมการแก้ไขปัญหาเสพติดในกลุ่มเยาวชนผู้มีโอกาสเข้าไปเกี่ยวข้องกับยาเสพติด. วิทยานิพนธ์สังคมสงเคราะห์ศาสตรมหาบัณฑิต, สาขาวิชาสังคมสงเคราะห์ศาสตร์, คณะสังคมสงเคราะห์ศาสตร์, มหาวิทยาลัยธรรมศาสตร์.
- ชญัญภัทก์ วงศ์บา. (2555). ประเมินผลการฝึกอบรมตามโครงการ Champion โดยประยุกต์ใช้แนวทางการประเมินของเคิร์กแพทริก. สำนักงานป้องกันควบคุมโรคที่ 6 ขอนแก่น, 19(1), 27-38.
- ชิดเชื้อ กนกเพชรรัตน์. (2550). รายงานการวิจัย การสร้างแบบทดสอบมาตรฐานวิชาเครื่องมือและการวัดอิเล็กทรอนิกส์ ตามหลักสูตรประกาศนียบัตรวิชาชีพชั้นสูง พุทธศักราช 2548 สถาบันเทคโนโลยีราชมงคล. กรุงเทพฯ: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- เขาว์ อินไย. (2553). การประเมินโครงการ. กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ไชยยศ ไพโรยศิริธรรม. (2550). การพัฒนารูปแบบการประเมินหลักสูตร: การประยุกต์ใช้การประเมินอภิमान. คุษฎีนิพนธ์การศึกษาคุษฎีบัณฑิต, สาขาวิชาการวิจัยและพัฒนาหลักสูตร, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- ทวีศักดิ์ อริยวัฒน์วงศ์. (2551). การประเมินโครงการพัฒนาวิษย์นักเรียนโรงเรียนวัดไร่ขิงวิทยา. สารนิพนธ์การศึกษามหาบัณฑิต, สาขาวิชาการบริหารการศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยศรีนครินทรวิโรฒ.
- นงลักษณ์ วิรัชชัย. (2551). การประเมินอภิमान และการวิเคราะห์อภิमान รายงานผลการประเมินคุณภาพภายนอกระดับอุดมศึกษา. วารสารวิธีวิทยาการวิจัย, 21(2), 125-145.



- เนติ เกลียวาเรศ. (2556). การประเมินอภิमान: แนวคิดและหลักการ. *วารสารวิชาการ*  
*เครือข่ายบัณฑิตศึกษามหาวิทยาลัยราชภัฏภาคเหนือ*, 3(4), 1-14.
- ปยุตา จันทร์เส็ง. (2557). การประเมินโครงการกองทุนพัฒนาบทบาทสตรีในเขตตำบลบ่อ  
 อำเภอลำดวน จังหวัดจันทบุรี. งานนิพนธ์รัฐประศาสนศาสตรมหาบัณฑิต, สาขาวิชา  
 การจัดการภาครัฐและภาคเอกชน, วิทยาลัยการบริหารรัฐกิจ, มหาวิทยาลัยบูรพา.
- ปิยวรรณ กระทบงาม. (2556). การประเมินอภิमानการประเมินคุณภาพภายนอกกรอบสาม  
 ของ สมศ. ในโรงเรียนระดับการศึกษาขั้นพื้นฐานจังหวัดชลบุรี. วิทยานิพนธ์  
 วิทยาศาสตร์มหาบัณฑิต, สาขาวิชาวิจัย วัดผลและสถิติการศึกษา, คณะศึกษาศาสตร์,  
 มหาวิทยาลัยบูรพา.
- พงศ์เทพ จิระโร. (2554). การวิจัยประเมิน โครงการ. ชลบุรี: บัณฑิตเอกสาร.
- พงศ์เทพ จิระโร. (2559). หลักการวิจัยทางการศึกษา (พิมพ์ครั้งที่ 7). ชลบุรี: บัณฑิตเอกสาร.
- พรพรรณ วันเอก. (2552). การสร้างแบบทดสอบมาตรฐานความรู้ด้านสหกรณ์สำหรับสมาชิก.  
 วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต, สาขาวิชาเศรษฐศาสตร์สหกรณ์, บัณฑิตวิทยาลัย,  
 มหาวิทยาลัยเกษตรศาสตร์.
- พิชิต ฤทธิจรูญ. (2555). เทคนิคการประเมินโครงการ. กรุงเทพฯ: แฮ็ลล์ออฟเคอร์มีส์.
- พิชญ์ ฟองสี. (2550). การประเมินทางการศึกษา: แนวคิดสู่การปฏิบัติ. กรุงเทพฯ: ด้านสุทธา  
 การพิมพ์.
- พิชญ์ ฟองสี. (2551). เทคนิคการประเมินโครงการ (พิมพ์ครั้งที่ 6). กรุงเทพฯ: ด้านสุทธาการพิมพ์.
- เพ็ญศรี เศรษฐวงค์. (2549). แบบทดสอบมาตรฐาน (Standardized test). *วารสารรามคำแหง*, 23(4),  
 373-392.
- มานะ ทองรักษ์. (2549). การบริหารจัดการโรงเรียนเพื่อพัฒนาสู่ความเป็นเลิศ. กรุงเทพฯ:  
 ธารอักษร.
- มาเรียม นิลพันธุ์. (2553). วิจัยทางการศึกษา (พิมพ์ครั้งที่ 2). นครปฐม: ศูนย์วิจัยและพัฒนา  
 ทางการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร.
- รัตนะ บัวสนธ์. (2550). ทิศทางและอาณาบริเวณการประเมิน. กรุงเทพฯ: สำนักพิมพ์แห่ง  
 จุฬาลงกรณ์มหาวิทยาลัย.
- รัตนะ บัวสนธ์. (2555). วิธีการเชิงผสมผสานสำหรับการวิจัยและการประเมิน. กรุงเทพฯ:  
 สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

- ศันสนีย์ อุตมอ่าง. (2554). รายงานการประเมินโครงการ การฝึกอบรมเชิงปฏิบัติการการผลิต  
ของม้วนเพื่อกระตุ้นเศรษฐกิจชุมชนตามแนวเศรษฐกิจพอเพียง. เพชรบูรณ์:  
คณะเทคโนโลยีการเกษตร มหาวิทยาลัยราชภัฏเพชรบูรณ์.
- ศิริชัย กาญจนวาสี. (2545). *ทฤษฎีการประเมิน* (พิมพ์ครั้งที่ 3). กรุงเทพฯ:  
สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. (2550). *ทฤษฎีการประเมิน* (พิมพ์ครั้งที่ 5). กรุงเทพฯ: สำนักพิมพ์แห่ง  
จุฬาลงกรณ์มหาวิทยาลัย.
- สมคิด พรหมจ้อย. (2550). *เทคนิคการประเมินโครงการ* (พิมพ์ครั้งที่ 5). กรุงเทพฯ: สำนักพิมพ์  
มหาวิทยาลัยสุโขทัยธรรมมาธิราช.
- สมนึก กัททัยชนิ. (2551). *การวัดผลการศึกษา* (พิมพ์ครั้งที่ 5). กทม. : ประสานการพิมพ์.
- สมหวัง พิธิยานุวัฒน์. (2549). *วิธีวิทยาการประเมินศาสตร์แห่งคุณค่า* (พิมพ์ครั้งที่ 3).  
กรุงเทพฯ: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สมหวัง พิธิยานุวัฒน์. (2553). *วิธีวิทยาการประเมิน: ศาสตร์แห่งคุณค่า* (พิมพ์ครั้งที่ 5). กรุงเทพฯ:  
จุฬาลงกรณ์มหาวิทยาลัย.
- สิรินธร สิ้นจินดาวงศ์. (2552). *การพัฒนาและประยุกต์ใช้มาตรฐานการประเมินอภิมาน  
สำหรับประเมินผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษาในประเทศไทย.  
คุณูปนิพนธ์ครุศาสตร์คุณูปบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา,  
บัณฑิตวิทยาลัย, จุฬาลงกรณ์มหาวิทยาลัย.*
- สิริศักดิ์ อัจฉิชัย. (2556). *การพัฒนารูปแบบการประเมินอภิมานสถานศึกษาของสำนักงาน  
เขตพื้นที่การศึกษา. วิทยานิพนธ์การศึกษาคุณูปบัณฑิต, สาขาวิชาวิจัยและประเมินผล  
การศึกษา, บัณฑิตวิทยาลัย, มหาวิทยาลัยนเรศวร.*
- สุวิมล ติรกันันท์. (2550). *การประเมินโครงการ: แนวทางสู่การปฏิบัติ* (พิมพ์ครั้งที่ 7). กรุงเทพฯ:  
สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- Kirkpatrick, D. L. (1975). *Techniques for evaluating, training program, evaluating training  
program*. San Francisco: American Society For Training and Development.
- Madaus, G. F., Scriven, M., & Stufflebeam, D. L. (1983). *Evaluation models*. Boston:  
KluwerNijhoff .
- Madaus, G. F., Scriven, M., & Stufflebeam, D. L. (1984). Educational evaluation and  
accountability: A review of quality assurance efforts. *The American*, 27(5), 649-673.
- Popham, W. J. (1993). *Educational evaluation*. Boston: Allyn and Bacon.

- Scriven, M. (1991). *Evaluation thesaurus* (4<sup>th</sup> ed.). Newbury Park, CA: Sage.
- Stufflebeam, D. L. (1967). The use and abuse of evaluation in Title III. *Theory Into Practice*, 6, 126-133.
- Stufflebeam, D. L. (2000). The CIPP model for evaluation. In D.L. Stufflebeam, G. F. Madaus, & T. Kellaghan, (Eds.). *Evaluation models* (2<sup>nd</sup> ed.). Boston: Kluwer Academic.
- Stufflebeam, D. L. Scriven, M., & Madaus, G. F. (1983). *Evaluation models*. Boston: Kluwer Nijhoff Publishing.
- Worthen, B. R., Sanders, J. R. (1987). *Educational evaluation: Alternative approaches and practical guidelines*. New York: Longman.

ภาคผนวก

ภาคผนวก ก  
รายชื่อผู้เชี่ยวชาญ

### รายชื่อผู้เชี่ยวชาญ

1. นายวีรพัฒน์ กุดแถลง (ศษ.ม.การวัดและประเมินผลการศึกษา)  
หัวหน้าฝ่ายวิชาการ (นักบริหารการศึกษา ระดับต้น)  
ส่วนส่งเสริมการจัดการศึกษาในระบบและนอกระบบ สำนักงานการศึกษาเมืองพัทยา
2. นางสาวนพวรรณ สิทธิประเสริฐ (ศษ.ม. การบริหารการศึกษา)  
นักวิชาการศึกษา ระดับปฏิบัติการ ฝ่ายวิชาการ  
ส่วนส่งเสริมการจัดการศึกษาในระบบและนอกระบบ สำนักงานการศึกษาเมืองพัทยา
3. ดร.อนุรักษ์ บุญพยนต์ (ปร.ด. ไทยศึกษา)  
ครูชำนาญการ โรงเรียนเมืองพัทยา 7 (บ้านหนองพังแค) สำนักงานการศึกษาเมืองพัทยา
4. ดร.ศจี จิระโร (ค.ศ. วัดผลและประเมินผลการศึกษา)  
อาจารย์ประจำภาควิชาการวัดและประเมินผลการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช
5. ดร.รุ่งนภา จันทรา (ปร.ด. การวิจัย วัดผลและสถิติการศึกษา)  
อาจารย์ประจำวิทยาลัยพยาบาลบรมราชชนนี

(สำเนา)

ที่ ศธ 6218/ว84

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา  
ต.แสนสุข อ.เมือง จ.ชลบุรี 20131

15 กรกฎาคม 2559

เรื่อง ขอความอนุเคราะห์ในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัย

เรียน (สำเนาดังแนบ)

สิ่งที่ส่งมาด้วย แก้วโครงย่อวิทยานิพนธ์ และเครื่องมือเพื่อการวิจัย จำนวน 1 ชุด

ด้วยนางสาวคำปิ่น ศรีมหาไชย นิสิตระดับบัณฑิตศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัตถุประสงค์ศึกษา มหาวิทยาลัยบูรพา ได้รับอนุมัติให้ทำวิทยานิพนธ์ เรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอกิมาน” ในความควบคุมดูแลของ ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร ประธานกรรมการ ขณะนี้อยู่ในขั้นตอนการสร้างเครื่องมือเพื่อการวิจัย ในการนี้ คณะศึกษาศาสตร์ได้พิจารณาแล้วเห็นว่าท่านเป็นผู้เชี่ยวชาญในเรื่องดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์จากท่านในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัยของนิสิตในครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา หวังเป็นอย่างยิ่งว่า คงได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ลงชื่อ)

เชษฐ ศิริสวัสดิ์

(ผู้ช่วยศาสตราจารย์ ดร.เชษฐ ศิริสวัสดิ์)

รองคณบดีฝ่ายบัณฑิตศึกษา ปฏิบัติราชการแทน

คณบดีคณะศึกษาศาสตร์ ปฏิบัติการแทน

ผู้รักษาการแทนอธิการบดีมหาวิทยาลัยบูรพา

ภาควิชาวิจัยและจิตวิทยาประยุกต์

โทรศัพท์ 0-3830-2076

โทรสาร 0-3839-7151

ผู้วิจัยโทร. 087-702-47734

(สำเนา)

ที่ ศธ 6218/ว84

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา  
ต.แสนสุข อ.เมือง จ.ชลบุรี 20131

15 กรกฎาคม 2559

เรื่อง ขอบความอนุเคราะห์ในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัย

เรียน นายวีรวัฒน์ กุดแถลง

สิ่งที่ส่งมาด้วย แก้วโครงย่อวิทยานิพนธ์ และเครื่องมือเพื่อการวิจัย จำนวน 1 ชุด

ด้วยนางสาวคำปิ่น ศรีมหาไชย นิสิตระดับบัณฑิตศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัตถุประสงค์ศึกษา มหาวิทยาลัยบูรพา ได้รับอนุมัติให้ทำวิทยานิพนธ์ เรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอกิमान” ในความควบคุมดูแลของ ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร ประธานกรรมการ ขณะนี้อยู่ในขั้นตอนการสร้างเครื่องมือเพื่อการวิจัย ในการนี้ คณะศึกษาศาสตร์ได้พิจารณาแล้วเห็นว่าท่านเป็นผู้เชี่ยวชาญในเรื่องดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์จากท่านในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัยของนิสิตในครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา หวังเป็นอย่างยิ่งว่า คงได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ลงชื่อ)

เชษฐ ศิริสวัสดิ์

(ผู้ช่วยศาสตราจารย์ ดร.เชษฐ ศิริสวัสดิ์)

รองคณบดีฝ่ายบัณฑิตศึกษา ปฏิบัติราชการแทน

คณบดีคณะศึกษาศาสตร์ ปฏิบัติการแทน

ผู้รักษาการแทนอธิการบดีมหาวิทยาลัยบูรพา

ภาควิชาวิจัยและจิตวิทยาประยุกต์

โทรศัพท์ 0-3830-2076

โทรสาร 0-3839-7151

ผู้วิจัยโทร. 087-702-47734



(สำเนา)

ที่ ศธ 6218/ว84

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา  
ต.แสนสุข อ.เมือง จ.ชลบุรี 20131

15 กรกฎาคม 2559

เรื่อง ขอบความอนุเคราะห์ในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัย

เรียน นางสาวนพวรรณ สิทธิประเสริฐ

สิ่งที่ส่งมาด้วย เค้าโครงย่อวิทยานิพนธ์ และเครื่องมือเพื่อการวิจัย จำนวน 1 ชุด

ด้วยนางสาวคำปิ่น ศรีมหาไชย นิสิตระดับบัณฑิตศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัตถุประสงค์ศึกษา มหาวิทยาลัยบูรพา ได้รับอนุมัติให้ทำวิทยานิพนธ์ เรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอกิมาน” ในความควบคุมดูแลของ ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร ประธานกรรมการ ขณะนี้อยู่ในขั้นตอนการสร้างเครื่องมือเพื่อการวิจัย ในการนี้ คณะศึกษาศาสตร์ได้พิจารณาแล้วเห็นว่าท่านเป็นผู้เชี่ยวชาญในเรื่องดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์จากท่านในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัยของนิสิตในครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา หวังเป็นอย่างยิ่งว่า คงได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ลงชื่อ)

เชษฐ ศิริสวัสดิ์

(ผู้ช่วยศาสตราจารย์ ดร.เชษฐ ศิริสวัสดิ์)

รองคณบดีฝ่ายบัณฑิตศึกษา ปฏิบัติราชการแทน

คณบดีคณะศึกษาศาสตร์ ปฏิบัติการแทน

ผู้รักษาการแทนอธิการบดีมหาวิทยาลัยบูรพา

ภาควิชาวิจัยและจิตวิทยาประยุกต์

โทรศัพท์ 0-3830-2076

โทรสาร 0-3839-7151

ผู้วิจัยโทร. 087-702-47734

(สำเนา)

ที่ ศธ 6218/ว84

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา  
ต.แสนสุข อ.เมือง จ.ชลบุรี 20131

15 กรกฎาคม 2559

เรื่อง ขอบความอนุเคราะห์ในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัย

เรียน ดร.อนุรักษ์ บุญพยนต์

สิ่งที่ส่งมาด้วย เค้าโครงย่อวิทยานิพนธ์ และเครื่องมือเพื่อการวิจัย จำนวน 1 ชุด

ด้วยนางสาวคำปิ่น ศรีมหาไชย นิสิตระดับบัณฑิตศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัตถุประสงค์ศึกษา มหาวิทยาลัยบูรพา ได้รับอนุมัติให้ทำวิทยานิพนธ์ เรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอกิमान” ในความควบคุมดูแลของ ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร ประธานกรรมการ ขณะนี้อยู่ในขั้นตอนการสร้างเครื่องมือเพื่อการวิจัย ในการนี้ คณะศึกษาศาสตร์ได้พิจารณาแล้วเห็นว่าท่านเป็นผู้เชี่ยวชาญในเรื่องดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์จากท่านในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัยของนิสิตในครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา หวังเป็นอย่างยิ่งว่า คงได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ลงชื่อ)

เชษฐ ศิริสวัสดิ์

(ผู้ช่วยศาสตราจารย์ ดร.เชษฐ ศิริสวัสดิ์)

รองคณบดีฝ่ายบัณฑิตศึกษา ปฏิบัติราชการแทน

คณบดีคณะศึกษาศาสตร์ ปฏิบัติการแทน

ผู้รักษาการแทนอธิการบดีมหาวิทยาลัยบูรพา

ภาควิชาวิจัยและจิตวิทยาประยุกต์

โทรศัพท์ 0-3830-2076

โทรสาร 0-3839-7151

ผู้วิจัยโทร. 087-702-47734

(สำเนา)

ที่ ศธ 6218/ว84

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา  
ต.แสนสุข อ.เมือง จ.ชลบุรี 20131

15 กรกฎาคม 2559

เรื่อง ขอบความอนุเคราะห์ในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัย

เรียน ดร.ศจี จิระโร

สิ่งที่ส่งมาด้วย เค้าโครงย่อวิทยานิพนธ์ และเครื่องมือเพื่อการวิจัย จำนวน 1 ชุด

ด้วยนางสาวคำปิ่น ศรีมหาไชย นิสิตระดับบัณฑิตศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัฒนและสถิติศึกษา มหาวิทยาลัยบูรพา ได้รับอนุมัติให้ทำวิทยานิพนธ์ เรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอกิमान” ในความควบคุมดูแลของ ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร ประธานกรรมการ ขณะนี้อยู่ในขั้นตอนการสร้างเครื่องมือเพื่อการวิจัย ในการนี้ คณะศึกษาศาสตร์ได้พิจารณาแล้วเห็นว่าท่านเป็นผู้เชี่ยวชาญในเรื่องดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์จากท่านในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัยของนิสิตในครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา หวังเป็นอย่างยิ่งว่า คงได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ลงชื่อ)

เชษฐ ศิริสวัสดิ์

(ผู้ช่วยศาสตราจารย์ ดร.เชษฐ ศิริสวัสดิ์)

รองคณบดีฝ่ายบัณฑิตศึกษา ปฏิบัติราชการแทน

คณบดีคณะศึกษาศาสตร์ ปฏิบัติการแทน

ผู้รักษาการแทนอธิการบดีมหาวิทยาลัยบูรพา

ภาควิชาวิจัยและจิตวิทยาประยุกต์

โทรศัพท์ 0-3830-2076

โทรสาร 0-3839-7151

ผู้วิจัยโทร. 087-702-47734

(สำเนา)

ที่ ศธ 6218/ว84

คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา  
ต.แสนสุข อ.เมือง จ.ชลบุรี 20131

15 กรกฎาคม 2559

เรื่อง ขอบความอนุเคราะห์ในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัย

เรียน ดร.รุ่งนภา จันทรา

สิ่งที่ส่งมาด้วย เค้าโครงย่อวิทยานิพนธ์ และเครื่องมือเพื่อการวิจัย จำนวน 1 ชุด

ด้วยนางสาวคำปิ่น ศรีมหาไชย นิสิตระดับบัณฑิตศึกษา หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิจัย วัตถุประสงค์ศึกษา มหาวิทยาลัยบูรพา ได้รับอนุมัติให้ทำวิทยานิพนธ์ เรื่อง “การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียนโรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมินอกิมาน” ในความควบคุมดูแลของ ผู้ช่วยศาสตราจารย์ นาวาตรี ดร.พงศ์เทพ จิระโร ประธานกรรมการ ขณะนี้อยู่ในขั้นตอนการสร้างเครื่องมือเพื่อการวิจัย ในการนี้ คณะศึกษาศาสตร์ได้พิจารณาแล้วเห็นว่าท่านเป็นผู้เชี่ยวชาญในเรื่องดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์จากท่านในการตรวจสอบความเที่ยงตรงของเครื่องมือเพื่อการวิจัยของนิสิตในครั้งนี้

จึงเรียนมาเพื่อโปรดพิจารณา คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา หวังเป็นอย่างยิ่งว่า คงได้รับความอนุเคราะห์จากท่านด้วยดี และขอขอบคุณอย่างสูงมา ณ โอกาสนี้

ขอแสดงความนับถือ

(ลงชื่อ)

เชษฐ ศิริสวัสดิ์

(ผู้ช่วยศาสตราจารย์ ดร.เชษฐ ศิริสวัสดิ์)

รองคณบดีฝ่ายบัณฑิตศึกษา ปฏิบัติราชการแทน

คณบดีคณะศึกษาศาสตร์ ปฏิบัติการแทน

ผู้รักษาการแทนอธิการบดีมหาวิทยาลัยบูรพา

ภาควิชาวิจัยและจิตวิทยาประยุกต์

โทรศัพท์ 0-3830-2076

โทรสาร 0-3839-7151

ผู้วิจัยโทร. 087-702-47734

ภาคผนวก ข  
เครื่องมือที่ใช้ในการวิจัย

### แบบประเมินโครงการ

“การประเมินโครงการสร้างข้อสอบมาตรฐานวัดความรู้ความสามารถในการคิดขั้นสูงของนักเรียน  
โรงเรียนในสังกัดเมืองพัทยา โดยบูรณาการรูปแบบการประเมินของเคิร์กแพตทริก  
และการประเมินอภิมาน”

#### คำชี้แจง

1. แบบประเมิน โครงการชุดนี้ มี 3 ตอน  
ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมิน  
ตอนที่ 2 การบูรณาการรูปแบบการประเมินของเคิร์กแพตทริกและการประเมิน

#### อภิมาน

- ตอนที่ 3 ข้อคิดเห็น/ ข้อเสนอแนะ
2. คำตอบของท่าน ผู้วิจัยจะนำไปใช้เพื่อประโยชน์ในการวิจัยเท่านั้น การนำเสนอ  
ผลวิจัยจะนำเสนอโดยภาพรวม ไม่มีผลกระทบต่อท่านเป็นส่วนตัวแต่อย่างใด

ขอขอบพระคุณเป็นอย่างสูงที่กรุณาตอบแบบสอบถามในครั้งนี้

นางสาวคำปิ่น ศรีมหาไชย  
นิสิตปริญญาโท สาขาวิจัย วัตถุประสงค์ และสถิติการศึกษา  
คณะศึกษาศาสตร์ มหาวิทยาลัยบูรพา

ตอนที่ 1 ข้อมูลทั่วไปของผู้ตอบแบบประเมิน

1. เพศ  ชาย  หญิง
2. อายุ.....ปี (มากกว่า 6 เดือนนับเป็น 1 ปี)
3. ประสบการณ์การสอน.....ปี (มากกว่า 6 เดือนนับเป็น 1 ปี)
4. ตำแหน่ง/ วิทยฐานะ
 

<input type="checkbox"/> ครูผู้ช่วย	<input type="checkbox"/> ครู	<input type="checkbox"/> ครูชำนาญการ
<input type="checkbox"/> ครูชำนาญการพิเศษ	<input type="checkbox"/> ครูเชี่ยวชาญ	<input type="checkbox"/> ครูเชี่ยวชาญพิเศษ
<input type="checkbox"/> อื่น ๆ (ระบุ) .....		
5. ระดับชั้นที่สอน
 

<input type="checkbox"/> ประถมต้น	<input type="checkbox"/> ประถมปลาย
<input type="checkbox"/> มัธยมศึกษาตอนต้น	<input type="checkbox"/> มัธยมศึกษาตอนปลาย
<input type="checkbox"/> อื่น ๆ (ระบุ) .....	
6. กลุ่มสาระการเรียนรู้ที่สอน
 

<input type="checkbox"/> ภาษาไทย	<input type="checkbox"/> คณิตศาสตร์
<input type="checkbox"/> วิทยาศาสตร์	<input type="checkbox"/> สังคมศึกษา ศาสนาและวัฒนธรรม
<input type="checkbox"/> สุขศึกษาและพลศึกษา	<input type="checkbox"/> ศิลปะศึกษา ดนตรี และนาฏศิลป์
<input type="checkbox"/> การงานพื้นฐานอาชีพและเทคโนโลยี	<input type="checkbox"/> ภาษาต่างประเทศ
<input type="checkbox"/> อื่น ๆ (ระบุ) .....	
7. วุฒิการศึกษาสูงสุด
 

<input type="checkbox"/> ปริญญาตรี	<input type="checkbox"/> ปริญญาโท
------------------------------------	-----------------------------------
8. วุฒิจำแนกตามสาขา
 

<input type="checkbox"/> วุฒิทางการศึกษา	<input type="checkbox"/> วุฒิสายาอื่น ๆ
--	---

ตอนที่ 2 การบูรณาการรูปแบบการประเมินของเครื่องแพดทริกและการประเมินอภิमान

รายการ	ระดับความคิดเห็น				
	มากที่สุด (5)	มาก (4)	ปานกลาง (3)	น้อย (2)	น้อยที่สุด (1)
<b>1. การประเมินปฏิริยาตอบสนอง</b>					
1.1 ความเหมาะสมของสถานที่ในการจัดโครงการ					
1.2 ความเหมาะสมของวัน-เวลาในการจัดโครงการ					
1.3 ความเหมาะสมของวิทยากร					
1.4 ความเหมาะสมของเอกสารประกอบการบรรยาย					
1.5 ความเหมาะสมหลักสูตรในแต่ละกิจกรรมกับ จำนวนเวลาในการฝึกอบรม					
1.6 ความเหมาะสมของสิ่งอำนวยความสะดวก					
1.7 ความเหมาะสมเพียงพอของสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในการประชุม					
1.8 ความเหมาะสมของชนิดโสตทัศนูปกรณ์					
1.9 ความเหมาะสมของความพร้อมอุปกรณ์ โสตทัศนูปกรณ์					
1.10 ความเหมาะสมการลำดับกิจกรรมในการ ฝึกอบรม					
<b>2. การประเมินการเรียนรู้</b>					
2.1 ประโยชน์ในด้านความรู้					
2.1.1 ความรู้ ความเข้าใจในเรื่องนี้ก่อนการอบรม					
2.1.2 ความรู้ ความเข้าใจในเรื่องนี้หลังการอบรม					
2.2 ประโยชน์ด้านความคิด					
2.2.1 สามารถจัดระบบความคิด					
2.2.2 สามารถประมวลความคิดสู่การพัฒนางาน อย่างเป็นระบบ					



รายการ	ระดับความคิดเห็น				
	มากที่สุด (5)	มาก (4)	ปานกลาง (3)	น้อย (2)	น้อยที่สุด (1)
2.3 ประโยชน์ด้านทักษะ/ การปฏิบัติ					
2.3.1 เกิดทักษะและประสบการณ์ใหม่ ๆ จากการฝึกอบรม					
2.3.2 สามารถปฏิบัติกรนอกแบบทดสอบวัดความรู้ความสามารถขั้นสูงของนักเรียนได้					
2.4 ประโยชน์ด้านทัศนคติ/ ค่านิยม					
2.4.1 ประโยชน์ที่ท่านได้รับจากโครงการ					
2.4.2 สิ่งที่ท่านได้รับจากโครงการนี้ตรงตามความคาดหวัง					
2.5 เนื้อหาที่บรรยายถูกต้องตรงกับหัวข้อที่กำหนด					
<b>3. การประเมินพฤติกรรม</b>					
3.1 สามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการปฏิบัติงานได้					
3.2 สามารถนำความรู้ไปเผยแพร่/ ถ่ายทอดแก่ชุมชนได้					
3.3 สามารถให้คำปรึกษาแก่เพื่อนร่วมงานได้					
3.4 มีความมั่นใจและสามารถนำความรู้ที่ได้รับไปใช้ได้					
3.5 สามารถแลกเปลี่ยนเรียนรู้กับเพื่อนร่วมงานได้					
<b>4. การประเมินผลลัพธ์ที่เกิดต่อองค์กร</b>					
4.1 ประโยชน์ที่ได้รับสามารถบูรณาการสู่การทำงานเป็นทีมได้					
4.2 ประโยชน์ที่ได้รับสามารถเพิ่มประสิทธิภาพการทำงาน					
4.3 ประโยชน์ที่ได้รับทำให้องค์กรยอมรับ					

รายการ	ระดับความคิดเห็น				
	มากที่สุด (5)	มาก (4)	ปานกลาง (3)	น้อย (2)	น้อยที่สุด (1)
4.4 ความเป็นไปได้ในการปฏิรูประบบการทำงานในการปฏิบัติงาน					
4.5 มีข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูงของนักเรียนในโรงเรียนสังกัดเมืองพัทยา					
4.6 ผลการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-net) ของนักเรียนในโรงเรียนสังกัดเมืองพัทยาสูงขึ้น					
4.7 ผลสัมฤทธิ์ทางการเรียนของนักเรียนในโรงเรียนสังกัดเมืองพัทยามีมาตรฐานที่ใกล้เคียงกัน					
4.8 มีคลังข้อสอบสำหรับพัฒนาผลสัมฤทธิ์ทางการเรียน					

ตอนที่ 3 ประโยชน์ที่ได้รับและข้อเสนอแนะ

ประโยชน์ที่ได้รับจากการเข้าร่วมอบรมในโครงการ

.....

.....

.....

ความเป็นไปได้หลังจากเข้าร่วมอบรมในโครงการ

.....

.....

.....

ความเหมาะสมของโครงการ

.....

.....

.....

ความถูกต้องด้านเนื้อหาที่ได้รับจากการเข้าร่วมอบรมโครงการ

.....  
.....  
.....

ข้อเสนอแนะ

.....  
.....  
.....

ภาคผนวก ค  
คุณภาพเครื่องมือที่ใช้ในการวิจัย

	รายการพิจารณา	คนที่				
		1	2	3	4	5
<b>1.</b>	<b>การประเมินปฏิริยาตอบสนอง</b>					
1.1	ความเหมาะสมของสถานที่ในการจัดโครงการ	1	1	1	1	1
1.2	ความเหมาะสมของวัน-เวลาในการจัดโครงการ	1	1	1	1	1
1.3	ความเหมาะสมของวิทยากร	1	1	1	1	1
1.4	ความเหมาะสมของเอกสารประกอบการบรรยาย	0	1	1	1	1
1.5	ความเหมาะสมหลักสูตรในแต่ละกิจกรรมกับจำนวนเวลาในการฝึกอบรม	1	1	1	1	1
1.6	ความเหมาะสมของสิ่งอำนวยความสะดวก	0	1	1	1	1
1.7	ความเหมาะสมเพียงพอของสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในการฝึกอบรม	0	1	1	1	1
1.8	ความเหมาะสมของชนิดวัสดุอุปกรณ์	1	1	1	1	1
1.9	ความเหมาะสมของความพร้อมอุปกรณ์วัสดุอุปกรณ์	1	1	1	1	1
1.10	ความเหมาะสมการลำดับกิจกรรมในการฝึกอบรม	1	1	1	1	1
<b>2.</b>	<b>การประเมินการเรียนรู้</b>					
2.1	ประโยชน์ในด้านความรู้					
2.1.1	ความรู้ ความเข้าใจในเรื่องนี้ก่อนการอบรม	1	1	1	1	1
2.1.2	ความรู้ ความเข้าใจในเรื่องนี้หลังการอบรม	1	1	1	1	1
2.2	ประโยชน์ด้านความคิด					
2.2.1	สามารถจัดระบบความคิด	1	1	1	1	1
2.2.2	สามารถประมวลความคิดสู่การพัฒนางานอย่างเป็นระบบ	1	1	1	1	1
2.3	ประโยชน์ด้านทักษะ/ การปฏิบัติ					
2.3.1	เกิดทักษะและประสบการณ์ใหม่ ๆ จากการฝึกอบรม	1	1	1	1	1
2.3.2	สามารถปฏิบัติการออกแบบทดสอบวัดความรู้ความสามารถขั้นสูงของนักเรียนได้	1	1	1	1	1

	รายการพิจารณา	คนที่				
		1	2	3	4	5
2.4	ประโยชน์ด้านทัศนคติ/ ค่านิยม					
2.4.1	ประโยชน์ที่ท่านได้รับจากโครงการ	1	1	1	1	1
2.4.2	สิ่งที่ท่านได้รับจากโครงการนี้ตรงตามความคาดหวัง	1	1	1	1	1
<b>3.</b>	<b>การประเมินพฤติกรรม</b>					
3.1	สามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ ในการปฏิบัติงานได้	1	1	1	1	1
3.2	สามารถนำความรู้ไปเผยแพร่/ ถ่ายทอด แก่เพื่อนร่วมงานได้	0	1	1	1	1
3.3	สามารถให้คำปรึกษาแก่เพื่อนร่วมงานได้	1	1	1	1	1
3.4	มีความมั่นใจและสามารถนำความรู้ที่ได้รับไปใช้ได้	1	1	1	1	1
3.5	สามารถแลกเปลี่ยนเรียนรู้กับเพื่อนร่วมงานได้	1	1	1	1	1
<b>4.</b>	<b>การประเมินผลลัพธ์ที่เกิดต่อองค์กร</b>					
4.1	ประโยชน์ที่ได้รับสามารถบูรณาการสู่การทำงาน เป็นทีมได้	1	1	1	1	1
4.2	ประโยชน์ที่ได้รับสามารถเพิ่มประสิทธิภาพการทำงาน	1	1	1	1	1
4.3	ประโยชน์ที่ได้รับทำให้องค์กรยอมรับ	1	1	1	1	1
4.4	ความเป็นไปได้ในการปฏิรูประบบการทำงาน ในการปฏิบัติงาน	1	1	1	1	1
4.5	มีข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้น สูงของนักเรียนในโรงเรียนสังกัดเมืองพัทยา	1	1	1	1	1
4.6	ผลการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-net) ของ นักเรียนใน โรงเรียนสังกัดเมืองพัทยา สูงขึ้น	1	1	1	1	1
4.7	ผลสัมฤทธิ์ทางการเรียนของนักเรียนใน โรงเรียนสังกัด เมืองพัทยามีมาตรฐานที่ใกล้เคียงกัน	1	1	1	1	1
4.8	มีคลังข้อสอบสำหรับพัฒนาผลสัมฤทธิ์ทางการเรียน	1	1	1	1	1

	รายการ	ดัชนีความสอดคล้อง
<b>1.</b>	<b>การประเมินปฏิริยาตอบสนอง</b>	
1.1	ความเหมาะสมของสถานที่ในการจัดโครงการ	1.00
1.2	ความเหมาะสมของวัน-เวลาในการจัดโครงการ	1.00
1.3	ความเหมาะสมของวิทยากร	1.00
1.4	ความเหมาะสมของเอกสารประกอบการบรรยาย	0.80
1.5	ความเหมาะสมหลักสูตรในแต่ละกิจกรรมกับจำนวนเวลาในการฝึกอบรม	1.00
1.6	ความเหมาะสมของสิ่งอำนวยความสะดวก	0.80
1.7	ความเหมาะสมเพียงพอของสิ่งอำนวยความสะดวกต่าง ๆ ที่ใช้ในฝึกอบรม	0.80
1.8	ความเหมาะสมของชนิดสื่อทัศนูปกรณ์	1.00
1.9	ความเหมาะสมของความพร้อมอุปกรณ์สื่อทัศนูปกรณ์	1.00
1.10	ความเหมาะสมการลำดับกิจกรรมในการฝึกอบรม	1.00
<b>2.</b>	<b>การประเมินการเรียนรู้</b>	
2.1	ประโยชน์ในด้านความรู้	
2.1.1	ความรู้ ความเข้าใจในเรื่องนี้ก่อนการอบรม	1.00
2.1.2	ความรู้ ความเข้าใจในเรื่องนี้หลังการอบรม	1.00
2.2	ประโยชน์ด้านความคิด	
2.2.1	สามารถจัดระบบความคิด	1.00
2.2.2	สามารถประมวลความคิดสู่การพัฒนางานอย่างเป็นระบบ	1.00
2.3	ประโยชน์ด้านทักษะ/ การปฏิบัติ	1.00
2.3.1	เกิดทักษะและประสบการณ์ใหม่ ๆ จากการฝึกอบรม	1.00
2.3.2	สามารถปฏิบัติการออกแบบทดสอบวัดความรู้ความสามารถขั้นสูงของนักเรียนได้	1.00
2.4	ประโยชน์ด้านทัศนคติ/ ค่านิยม	1.00
2.4.1	ประโยชน์ที่ท่านได้รับจากโครงการ	1.00
2.4.2	สิ่งที่ท่านได้รับจากโครงการนี้ตรงตามความคาดหวัง	1.00

	รายการ	ดัชนีความสอดคล้อง
<b>3.</b>	<b>การประเมินพฤติกรรม</b>	
3.1	สามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการปฏิบัติงานได้	1.00
3.2	สามารถนำความรู้ไปเผยแพร่/ ถ่ายทอดแก่เพื่อนร่วมงานได้	0.80
3.3	สามารถให้คำปรึกษาแก่เพื่อนร่วมงานได้	1.00
3.4	มีความมั่นใจและสามารถนำความรู้ที่ได้รับไปใช้ได้	1.00
3.5	สามารถแลกเปลี่ยนเรียนรู้กับเพื่อนร่วมงานได้	1.00
<b>4.</b>	<b>การประเมินผลลัพธ์ที่เกิดต่อองค์กร</b>	1.00
4.1	ประโยชน์ที่ได้รับสามารถบูรณาการสู่การทำงานเป็นทีมได้	1.00
4.2	ประโยชน์ที่ได้รับสามารถเพิ่มประสิทธิภาพการทำงาน	1.00
4.3	ประโยชน์ที่ได้รับทำให้องค์กรยอมรับ	1.00
4.4	ความเป็นไปได้ในการปฏิรูประบบการทำงานในการปฏิบัติงาน	1.00
4.5	มีข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูงของนักเรียนในโรงเรียนสังกัดเมืองพัทยา	1.00
4.6	ผลการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-Net) ของนักเรียนในโรงเรียนสังกัดเมืองพัทยาสูงขึ้น	1.00
4.7	ผลสัมฤทธิ์ทางการเรียนของนักเรียนในโรงเรียนสังกัดเมืองพัทยา มีมาตรฐานที่ใกล้เคียงกัน	1.00
4.8	มีคลังข้อสอบสำหรับพัฒนาผลสัมฤทธิ์ทางการเรียน	1.00

IOC = 0.80-1.00



ภาคผนวก ง

โครงการสร้างข้อสอบมาตรฐานวัดความรู้ ความสามารถในการคิดขั้นสูง

## โครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ ความสามารถในการคิดขั้นสูง

### หลักการและเหตุผล

การจัดการศึกษาจำเป็นต้องมีจุดมุ่งหมายที่ชัดเจน เพื่อช่วยกำหนดคุณลักษณะของผู้สำเร็จการศึกษาที่พึงประสงค์หลังจากได้รับการเรียนการสอน จุดมุ่งหมายยังช่วยให้ทิศทางของการจัดประสบการณ์การเรียนรู้หรือกิจกรรมการเรียนการสอนที่เหมาะสม การวัดและประเมินผลเข้ามามีบทบาทสำคัญในการตัดสินผลการเรียนรู้ของผู้เรียนว่าบรรลุผลสำเร็จตามจุดมุ่งหมายที่ต้องการหรือไม่ มากน้อยเพียงไร รวมทั้งช่วยให้สารสนเทศอันเป็นประโยชน์ต่อการปรับปรุงและพัฒนาการเรียนการสอน

การประเมินผลเป็นการตัดสินคุณค่าของสิ่งใดสิ่งหนึ่งด้วยวิธีการสืบสวนสอบสวนอย่างมีระบบแบบแผนการประเมินผลทางการศึกษาเป็นการตัดสินคุณค่าของสิ่งที่เกี่ยวข้องกับการประเมินผู้เรียนเพื่อตีคุณค่าของการเรียนรู้ที่เกิดขึ้นกับผู้เรียนแต่ละคนในการประเมินผล การเรียนรู้ผู้สอนสามารถกระทำได้ทุกขั้นตอนของการเรียนการสอนตั้งแต่ก่อนเริ่มเรียน ระหว่างการเรียนและหลังการเรียน โดยอาจใช้การประเมินทำหน้าที่ในการประเมินความก้าวหน้าเพื่อทราบพัฒนาการและปรับปรุงการเรียนรู้ หรือประเมินรวมสรุปเพื่อการตัดสินคุณค่าของการเรียนรู้รวมยอดของผู้เรียน

แบบทดสอบเป็นเครื่องมือในการวัดผลการศึกษาชนิดหนึ่งที่ใช้เป็นหลักในการวัดความสามารถต่าง ๆ ของผู้เรียนในทุกระดับการศึกษา การวัดผลด้วยแบบทดสอบ หรือการทดสอบนั้นจะถูกต้องเที่ยงตรงหรือเชื่อมั่นได้เพียงใดขึ้นอยู่กับคุณภาพของแบบทดสอบ ดังนั้น การเขียนข้อสอบที่ถูกต้องตามหลักเกณฑ์จึงเป็นสิ่งสมควร คำนึงถึงเป็นอย่างยิ่งเพราะจะส่งผลต่อคุณภาพของแบบทดสอบโดยตรง

การเขียนข้อสอบให้มีคุณภาพดีต้องอาศัยองค์ประกอบหลายประการ เช่น การวางแผน การกำหนดจุดมุ่งหมายการกำหนดคุณลักษณะที่จะวัด แต่สิ่งที่สำคัญที่สุดที่จะทำให้ข้อสอบดีได้นั้นคือ ตัวผู้เขียนข้อสอบมีลักษณะและคุณสมบัติอย่างไร บุคคลที่จะเขียนข้อสอบได้ดีควรมีลักษณะที่สำคัญ ดังนี้

1. มีความรู้ในเนื้อหา ถ้าขาดความรู้แม้จะเปิดหนังสือประกอบก็ตามย่อมเขียนข้อสอบให้ไม่ได้ยาก
2. รู้จุดมุ่งหมายของรายวิชา หรือรู้ลักษณะความสามารถที่ต้องการจะวัดนั้นเป็นอย่างดี เพื่อจะได้เขียนข้อสอบให้สามารถวัดได้ตรงตามจุดมุ่งหมายและลักษณะความสามารถนั้น ๆ

3. รู้วิธีการถาม คือ รู้เทคนิคและวิธีการตั้งคำถามเพื่อให้สามารถเลือกเฟ้นชนิดของคำถามที่เหมาะสม พฤติกรรมนั้นควรถามแบบไหน ด้วยคำถามประเภทใดจึงจะเหมาะสม และสามารถถามได้อย่างมีประสิทธิภาพ

4. มีทักษะในการใช้ภาษา เป็นคุณสมบัติที่จะช่วยให้ข้อสอบมีความชัดเจนจนเกิดความเป็นปรนัย เพราะคำถามที่ชัดเจน รัดกุม ใช้ภาษาได้อย่างเหมาะสมนั้นจะช่วยสื่อความหมายให้ผู้สอบเข้าใจความต้องการของผู้ถามได้อย่างตรงจุด

5. มีทักษะในการเขียนและวิจารณ์ข้อสอบ ถ้าเคยเขียนข้อสอบมากก็ย่อมเขียนข้อสอบได้คล่อง ยิ่งเคยพบเห็นข้อสอบมากก็ยิ่งทำให้มองแง่มุมการถามได้กว้างขวางขึ้น

จากความสำคัญดังกล่าวจะเห็นว่าทักษะการสร้างข้อสอบของครูเป็นสิ่งสำคัญที่จะมีส่วนช่วยในการพัฒนาผลสัมฤทธิ์ทางการเรียนของนักเรียนทั้งการทดสอบในระดับชั้นเรียน การทดสอบในระดับท้องถิ่น หรือแม้กระทั่งการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-Net) ดังนั้นสำนักงานการศึกษา โดยฝ่ายวิชาการ จึงได้จัดทำโครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูงขึ้น เพื่อพัฒนาความสามารถ ทักษะในการสร้างข้อสอบของพนักงานครูและพัฒนาผลสัมฤทธิ์ทางการเรียนของนักเรียนสังกัดเมืองพัทยาตามมาตรฐานการศึกษาต่อไป

### วัตถุประสงค์

1. เพื่อจัดทำข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูงของนักเรียนสังกัดเมืองพัทยา
2. เพื่อพัฒนาความสามารถของพนักงานครูสังกัดเมืองพัทยาในการพัฒนาข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูง

### เป้าหมาย

1. จัดอบรมพนักงานครูในการสร้างข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูงจำนวน 120 คน โดยอบรมทั้งหมด 4 ครั้ง
2. มีข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูงในทุกกลุ่มสาระการเรียนรู้

### ขั้นตอนการดำเนินงาน

#### ขั้นเตรียมการ

1. คัดเลือกคณะทำงานในการจัดทำข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูง ของนักเรียนใน โรงเรียนสังกัดเมืองพัทยา

2. คณะทำงานในการจัดทำข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูง  
ของนักเรียนในโรงเรียนสังกัดเมืองพัทยา จำนวน 114 คน ประกอบด้วย

หัวหน้าฝ่ายวิชาการ	1 คน
นักวิชาการศึกษา/ เจ้าหน้าที่โครงการ	3 คน
ผู้อำนวยการสถานศึกษา/ รองผู้อำนวยการสถานศึกษา (รับผิดชอบงานวิชาการ)	10 คน
ผู้แทนครูวัดผลทุกสถานศึกษา	11 คน
ครูผู้สอนกลุ่มสาระการเรียนรู้	89 คน

**ขั้นตอนดำเนินการ (ปีงบประมาณ 2559)**

เป็นการอบรมเชิงปฏิบัติการจำนวน 4 ครั้ง ครั้งละ 3 วัน วันละ 6 ชั่วโมง รวมเป็น  
72 ชั่วโมง รายละเอียดตามตาราง

ครั้งที่/ เวลา	กิจกรรม/ เนื้อหา	ผลลัพธ์
ครั้งที่ 1 วันที่ 27-29 เมษายน พ.ศ. 2559	1. ความรู้เบื้องต้นเกี่ยวกับการจัดทำ ข้อสอบมาตรฐานเพื่อวัดความคิดขั้นสูง 2. การจัดทำ test blue print	Test blueprint
ครั้งที่ 2 วันที่ 9-11 มิถุนายน พ.ศ. 2559	1. หลักการออกข้อสอบมาตรฐานเพื่อ วัดความคิดขั้นสูง 2. ฝึกปฏิบัติการออกข้อสอบมาตรฐาน เพื่อวัดความคิดขั้นสูง	ความรู้และทักษะของครู ในการออกข้อสอบมาตรฐาน เพื่อวัดความคิดขั้นสูง
ครั้งที่ 3 วันที่ 26-28 มิถุนายน พ.ศ. 2559	1. ทบทวนหลักการออกข้อสอบ มาตรฐานเพื่อวัดความคิดขั้นสูง 2. ฝึกปฏิบัติการออกข้อสอบมาตรฐาน เพื่อวัดความคิดขั้นสูง	ข้อสอบตัวชี้วัดละ 6-10ข้อ/ กลุ่มสาระการเรียนรู้/ ชั้นเรียน
ครั้งที่ 4 31 กรกฎาคม- 2 สิงหาคม พ.ศ. 2559	1. หลักการหาคุณภาพของข้อสอบ 2. ฝึกปฏิบัติการหาคุณภาพของ ข้อสอบ 3. การจัดชุดข้อสอบ	1. ความสามารถในการหา คุณภาพของข้อสอบ 2. ชุดข้อสอบทดลอง 1 ชุด/ กลุ่มสาระ/ ชั้น จำนวน 20 ข้อ 3. ชุดข้อสอบที่จะนำไป ทดลอง กลุ่มสาระฯ ละ 2 ชุด

### ขั้นการสรุปผล/ รายงาน

- นำข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูงไปใช้ทดสอบกับนักเรียนปีการศึกษา พ.ศ. 2559
- ตรวจ สรุป รายงานผลคะแนนให้สถานศึกษา และผู้บริหารเมืองพัทยา

### ระยะเวลาและสถานที่

ครั้งที่ 1 วันที่ 27-29 เมษายน พ.ศ. 2559	โรงเรียนเมืองพัทยา 1
ครั้งที่ 2 วันที่ 9-11 มิถุนายน พ.ศ. 2559	โรงเรียนเมืองพัทยา 1
ครั้งที่ 3 วันที่ 26-28 มิถุนายน พ.ศ. 2559	โรงเรียนเมืองพัทยา 1
ครั้งที่ 4 วันที่ 31 กรกฎาคม-2 สิงหาคม พ.ศ. 2559	โรงเรียนเมืองพัทยา 1

### งบประมาณ

ใช้งบประมาณจาก ฝ่ายวิชาการ งานบริหารทั่วไปเกี่ยวกับการศึกษาหมวดค่าใช้สอยประเภท รายจ่ายเกี่ยวกับการปฏิบัติราชการไม่เข้าลักษณะรายจ่ายหมวดอื่น โครงการสร้างข้อสอบมาตรฐานเพื่อวัดความรู้ความสามารถในการคิดขั้นสูง ตั้งไว้ 500,000 บาท (ห้าแสนบาท)

### หน่วยงานที่รับผิดชอบ

ฝ่ายวิชาการ สำนักการศึกษา

### ผลที่คาดว่าจะได้รับ

- ข้อสอบมาตรฐานเพื่อวัดความสามารถในการคิดขั้นสูง ของนักเรียนในโรงเรียนสังกัดเมืองพัทยา
- ผลการทดสอบทางการศึกษาขั้นพื้นฐานระดับชาติ (O-Net) ของนักเรียนในโรงเรียนสังกัดเมืองพัทยาสูงขึ้น
- ผลสัมฤทธิ์ทางการเรียนของนักเรียนใน โรงเรียนสังกัดเมืองพัทยา มีมาตรฐานที่ใกล้เคียงกัน
- มีคลังข้อสอบสำหรับพัฒนาผลสัมฤทธิ์ทางการเรียน