

การพัฒนาและประยุกต์ใช้มาตรฐานการประเมินอภิमान สำหรับประเมินรายงาน
ผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษาในประเทศไทย

นางสิรินธร สิ้นจินดาวงศ์

ศูนย์วิทยทรัพยากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต

สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2552

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

DEVELOPMENT AND APPLICATION OF THE META-EVALUATION STANDARDS TO
EVALUATE REPORT RESULTS OF INTERNAL QUALITY ASSESSMENT FOR HIGHER
EDUCATION INSTITUTIONS IN THAILAND

Mrs. Sirinthorn Sinjindawong

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Educational Measurement and Evaluation

Department of Educational Research and Psychology

Faculty of Education
Chulalongkorn University

Academic Year 2009

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์

การพัฒนาและประยุกต์ใช้มาตรฐานการประเมินอภิमान
สำหรับประเมินรายงานผลการประเมินคุณภาพภายในของ
สถาบันอุดมศึกษา ในประเทศไทย

โดย

นางสิรินธร สิ้นจินดาวงศ์

สาขาวิชา

การวัดและประเมินผลการศึกษา

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ผู้ช่วยศาสตราจารย์ ดร.ณัฐสุภรณ์ หลาวทอง

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี

คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัย
ของ การศึกษาตามหลักสูตรปริญญาศึกษาศาสตรบัณฑิต

..... คณบดีคณะครุศาสตร์
(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(ศาสตราจารย์กิตติคุณ ดร.นงลักษณ์ วิรัชชัย)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.ณัฐสุภรณ์ หลาวทอง)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี)

..... กรรมการ
(รองศาสตราจารย์ ดร. โชติกา ภาณีผล)

..... กรรมการภายนอกมหาวิทยาลัย
(ดร.คมศร วงษ์รักษา)

สิรินธร สิ้นจินดาวงศ์: การพัฒนาและประยุกต์ใช้มาตรฐานการประเมินอภีมาณสำหรับประเมิน รายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษาในประเทศไทย (DEVELOPMENT AND APPLICATION OF THE META-EVALUATION STANDARDS TO EVALUATE REPORT RESULTS OF INTERNAL QUALITY ASSESSMENT FOR HIGHER EDUCATION INSTITUTIONS IN THAILAND) อ. ที่ปรึกษาวิทยานิพนธ์หลัก: ผู้ช่วยศาสตราจารย์ ดร.ณัฐภรณ์ หลาวทอง, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม: ศาสตราจารย์ ดร.ศิริชัย กาญจนวาสี, 283 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ 1) พัฒนามาตรฐานการประเมินอภีมาณ สำหรับประเมินรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาในประเทศไทย 2) ตรวจสอบคุณภาพของมาตรฐานการประเมินอภีมาณ ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา 3) ประยุกต์ใช้มาตรฐานการประเมินอภีมาณในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา แหล่งข้อมูลในการวิจัย คือ รายงานผลการประเมินคุณภาพภายใน และรายงานการประเมินตนเองของสถาบันอุดมศึกษา ประจำปีการศึกษา 2550 จำนวน 50 ฉบับ และนักประเมินอภีมาณที่เข้ารับการอบรมจำนวน 2 รุ่น รวม 28 คน เครื่องมือที่ใช้ในการวิจัย คือ แบบประเมินอภีมาณ คู่มือการประเมินอภีมาณ และหลักสูตรอบรมพัฒนานักประเมินอภีมาณ วิเคราะห์ข้อมูลด้วยสถิติพื้นฐาน ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน สถิติทดสอบที การวิเคราะห์องค์ประกอบเชิงยืนยันและสัมประสิทธิ์การสรุปอ้างอิง ผลการวิจัยพบว่า

1. มาตรฐานการประเมินอภีมาณ ประกอบด้วย 5 มาตรฐาน คือ ความตรงของการประเมิน อรรถประโยชน์ จริยธรรมในการประเมิน ความน่าเชื่อถือของนักประเมิน และประสิทธิผลต้นทุน รวม 38 ตัวบ่งชี้
2. คุณภาพของมาตรฐานการประเมินอภีมาณ มีความตรงเชิงโครงสร้าง ด้วยเทคนิคกลุ่มผู้ชี้วัด โดยสามารถจำแนกคุณภาพรายงานผลการประเมินคุณภาพภายใน ออกเป็นกลุ่มสูง และกลุ่มต่ำ และความเที่ยงแบบสัมประสิทธิ์การสรุปอ้างอิง เมื่อใช้นักประเมินอภีมาณ 1, 2, 3 คน ประเมินรายงานผลการประเมินคุณภาพภายใน จำนวน 3 ฉบับ มีค่าเท่ากับ 0.542, 0.689, และ 0.758 ตามลำดับ หากประเมินรายงานฯ 5 ฉบับ มีค่าเท่ากับ 0.663, 0.787, และ 0.839 ตามลำดับ ประเมินรายงาน 7 ฉบับ มีค่าเท่ากับ 0.734, 0.838, และ 0.880 ตามลำดับ และประเมินรายงานฯ 9 ฉบับ มีค่าเท่ากับ 0.780, 0.869, และ 0.904 ตามลำดับ
3. ผลการประเมินอภีมาณของรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษาอยู่ในระดับดี โดยมาตรฐานที่มีผลการประเมินในระดับดี คือ ความน่าเชื่อถือของนักประเมิน อรรถประโยชน์ จริยธรรมในการประเมิน และความตรงของการประเมิน ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้

ภาควิชา วิจัยและจิตวิทยาการศึกษาลายมือชื่อนิสิต.....
 สาขาวิชา การวัดและประเมินผลการศึกษาลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
 ปีการศึกษา 2552ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....

4984739027 : MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION

KEYWORDS : META-EVALUATION STANDARDS / META-EVALUATOR / QUALITY ASSESSMENT /HIGHER EDUCATION

SIRINTHORN SINJINDAWONG: DEVELOPMENT AND APPLICATION OF THE META-EVALUATION STANDARDS TO EVALUATE REPORT RESULTS OF INTERNAL QUALITY ASSESSMENT FOR HIGHER EDUCATION INSTITUTIONS IN THAILAND. THESIS ADVISOR: ASST. PROF. NUTTAPORN LAWTHONG, Ph.D., THESIS CO-ADVISOR: PROF. SIRICHAJ KANJANAWASEE, Ph.D., 283 pp.

The purposes of this research were: 1) to develop the meta-evaluation standards for an evaluation of internal quality assessment in Thai higher education institutions. 2) to validate the meta-evaluation standards for evaluating report results of internal quality assessment for higher education institutions and 3) to apply the meta-evaluation standards to evaluate report results of internal quality assessment for higher education institutions. The data were collected from 50 internal quality assessment reports and self - assessment reports in higher education institutions, and 28 meta-evaluators. Three kinds of instruments consist of meta-evaluation checklist, meta-evaluation manual, and meta-evaluator training curriculum. The data analyses were conducted by statistics of quantitative data as: frequency, percentage, mean, standard deviation, t-test, confirmatory factor analysis, and generalizability coefficient.

The results were:

1. The new meta-evaluation standards consisted of 5 standards and 38 indicators: 1) validity 2) utility 3) ethicality 4) credibility and 5) cost-effectiveness.

2. The quality of meta-evaluation standard, the construct validity, used known-group technique which discriminated high and low quality reports. Reliability found that when using 1, 2, and 3 meta-evaluators evaluated per report, and g-coefficient from 3 reports were 0.542, 0.689, and 0.758 respectively. They also evaluated per report from 5, 7, and 9 reports, and g-coefficient from 5 reports were 0.663, 0.787, and 0.839 respectively, g-coefficient from 7 reports were 0.734, 0.838, and 0.880 respectively, and g-coefficient from 9 reports were 0.780, 0.869, and 0.904 respectively.

3. The quality of internal quality assessment reports in higher education institutions based on meta-evaluation standards were good level reports, which credibility, utility, ethicality, and validity were good, but cost-effectiveness was fair level.

Department :	Educational Research and Psychology	Student's Signature	<i>Sirinthorn S.</i>
Field of Study :	Educational Measurement and Evaluation	Advisor's Signature	<i>Nuttaporn L.</i>
Academic Year : 2009	Co-Advisor's Signature	<i>S. Kanjanawasee</i>

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี ด้วยความเอาใจใส่และปรารถนาดีจาก ผู้ช่วยศาสตราจารย์ ดร. ภูมิรัฐภรณ์ หลาวทอง อาจารย์ที่ปรึกษาวิทยานิพนธ์ และศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ซึ่งได้สละเวลาให้คำปรึกษาแนะนำผู้วิจัยเป็นอย่างดี ตลอดระยะเวลาการทำวิจัย รวมทั้งศาสตราจารย์กิตติคุณ ดร.นงลักษณ์ วิรัชชัย ประธานคณะกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร. ชาติกา ภาณุผล และ ดร.คมศร วงษ์รักษา คณะกรรมการสอบวิทยานิพนธ์ รวมถึงผู้ช่วยศาสตราจารย์ ดร.กมลวรรณ ตั้งธนานนท์ กรรมการสอบโครงร่างวิทยานิพนธ์ ที่ให้คำแนะนำในการทำวิทยานิพนธ์ และปรับปรุงแก้ไขจนสมบูรณ์ ผู้วิจัยรู้สึกซาบซึ้งและขอกราบขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้

ขอกราบขอบพระคุณ ศาสตราจารย์ ดร.สุวิมล ว่องวาณิช และดร.วราภรณ์ ไทยมา ที่คอยให้กำลังใจ ห่วงใย และให้ความช่วยเหลือผู้วิจัยมาโดยตลอด และขอขอบคุณผู้เชี่ยวชาญในการตรวจคุณภาพเครื่องมือที่ให้คำแนะนำในการแก้ไขปรับปรุงเครื่องมือในการวิจัย รวมถึงผู้ทรงคุณวุฒิด้านการประกันคุณภาพ ที่ทำการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา เพื่อให้งานวิจัยมีคุณภาพและสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ คุณอรสา ภาววิมล สำนักงานคณะกรรมการการอุดมศึกษาที่อำนวยความสะดวกให้ข้อมูล SAR และรายงานฯ ของสถาบันอุดมศึกษา และนอกจากนี้ขอขอบคุณนักประเมินอภิमानทุกท่าน ได้แก่ น.อ.พงษ์สุวรรณ ศรีสุวรรณ คุณศักดิ์สิทธิ์ ฤทธิรัตน์ คุณดวงเนตร ธรรมกุล คุณวราพร เจริญวรรณ คุณภัทราวดี มากมี คุณสาธิตา สฤตรัตน์กุลชัย คุณพนัส จันทร์เปล่ง คุณยุทธพงษ์ อายุสุข คุณทัศนศิริรินทร์ สว่างบุญ คุณจันทร์แรม แซ่เต็ง คุณมัยดี แวดราแม คุณมนัสนันท์ รอดเชื้อจิ้น อ. ฐิติภูมิ ว่องทรัพย์ทวี อ.ชวลิต มณีศรี ผศ. ชนิษฐา ชัยรัตนาวรรณ และผู้ประเมินคุณภาพภายในที่เข้าร่วมอบรมตามหลักสูตร รวมทั้งเสียสละเวลาอันมีค่าในการประเมินอภิमान จนสามารถสำเร็จตามเป้าหมาย รวมทั้งให้ข้อเสนอแนะในการประเมินที่เป็นประโยชน์เสมอมา

ขอขอบคุณจุฬาลงกรณ์มหาวิทยาลัยที่ให้โอกาสผู้วิจัยได้รับการสนับสนุนทุนวิจัยจาก “ทุน 90 ปี จุฬาลงกรณ์มหาวิทยาลัย” กองทุนรัชดาภิเษกสมโภช เพื่ออุดหนุนในการทำวิทยานิพนธ์จนสำเร็จลุล่วงไปด้วยดี

ขอกราบขอบพระคุณ มหาวิทยาลัยศรีปทุมที่ให้ทุนการศึกษา และเวลาในการศึกษาจนสำเร็จ การศึกษา และขอขอบคุณผู้อำนวยการสำนักงานแผนและประกันคุณภาพ รวมถึงพี่ๆ น้องๆ ในสำนักงานที่ช่วยประสานงานดำเนินการจัดอบรม และงานอื่นๆ จนสำเร็จลุล่วงไปด้วยดี

นอกจากนี้ ขอขอบคุณทุกท่านที่มีส่วนร่วม ให้วิทยานิพนธ์นี้สำเร็จตามเป้าหมาย ได้แก่ คุณพุดพงษ์ สุขสว่าง คุณชนะศึก นิชานนท์ ที่ให้คำปรึกษาในการวิเคราะห์ข้อมูล รวมถึงเพื่อนๆ พี่ๆ น้องๆ ที่ร่วมเรียนในภาควิชาวิจัยและจิตวิทยาการศึกษาทุกท่าน อีกทั้งผู้ที่ยกย่องอำนวยความสะดวก ได้แก่ คุณอุทุมพร ชาติเผือก คุณจอง หลายชั้น และท่านอื่นๆ ไม่ได้กล่าวนามไว้ ณ ที่นี้ด้วย

คุณค่า และประโยชน์ของวิทยานิพนธ์นี้ ผู้วิจัยขอมอบให้แก่ ผู้ที่อยู่เบื้องหลัง ความสำเร็จในครั้งนี้ คือ คุณพ่อไพโรจน์ คุณแม่เฉลีย์ สามีคุณพรชัย และลูกสาวเด็กหญิงนันทชา รวมทั้งญาติพี่น้องทุกท่าน ที่คอยให้การสนับสนุน เป็นกำลังใจให้ความรัก และความห่วงใยต่อผู้วิจัยตลอดมา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ฅ
สารบัญภาพ.....	ฎ
บทที่	
1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
คำถามการวิจัย.....	8
วัตถุประสงค์ของการวิจัย.....	8
ขอบเขตของการวิจัย.....	8
คำจำกัดความที่ใช้ในการวิจัย.....	9
ประโยชน์ที่ได้รับ.....	11
2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	13
ตอนที่ 1 แนวคิดทฤษฎีเกี่ยวกับการประกันคุณภาพ.....	13
ตอนที่ 2 มาตรฐานการประเมิน เกณฑ์การประเมิน.....	24
ตอนที่ 3 แนวคิดทฤษฎีเกี่ยวกับการประเมินอภิมาน.....	35
ตอนที่ 4 แนวทางการเขียนรายงานการประเมิน.....	59
ตอนที่ 5 กรอบแนวคิดในการวิจัย.....	67
3 วิธีดำเนินการวิจัย.....	71
ขั้นตอนการดำเนินการวิจัย.....	71
แหล่งข้อมูลสำหรับการวิจัย.....	75
เครื่องมือที่ใช้ในการวิจัย.....	76
การเก็บรวบรวมข้อมูล.....	85
การวิเคราะห์ข้อมูล.....	88
สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	89

บทที่	หน้า
4 ผลการวิเคราะห์ข้อมูล.....	92
ผลการวิเคราะห์.....	93
ตอนที่ 1 การพัฒนา และวิเคราะห์มาตรฐานการประเมินอภิमान.....	93
ตอนที่ 2 ผลการอบรมพัฒนานักประเมินอภิमान.....	121
ตอนที่ 3 ผลการตรวจสอบคุณภาพของเครื่องมือประเมินอภิमान.....	129
ตอนที่ 4 ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของ สถาบันอุดมศึกษา.....	133
5 สรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ.....	162
สรุปผลการวิจัย.....	163
อภิปรายผลการวิจัย.....	171
ข้อเสนอแนะ.....	179
รายการอ้างอิง.....	182
ภาคผนวก.....	191
ภาคผนวก ก รายนามผู้เชี่ยวชาญ และผู้ทรงคุณวุฒิ.....	192
ภาคผนวก ข เครื่องมือในการวิจัย.....	195
ภาคผนวก ค ภาพกิจกรรมการอบรมนักประเมินอภิमान.....	244
ภาคผนวก ง Print Out การวิเคราะห์ข้อมูล.....	247
ประวัติผู้เขียนวิทยานิพนธ์.....	283

สารบัญภาพ

ภาพที่	หน้า
2.1	15
2.2	21
2.3	68
2.4	69
2.5	70
3.1	74
4.1	109
4.2	111
4.3	114
4.4	116
4.5	118
4.6	120
4.7	120
4.8	130
4.9	150
4.10	150
4.11	150
4.12	150
4.13	150

สารบัญตาราง

ตารางที่	หน้า
2.1 ความสัมพันธ์ระหว่างมาตรฐานการศึกษาชาติ กับการประเมินคุณภาพ การศึกษา.....	20
2.2 การเปรียบเทียบการประยุกต์ใช้ The Program Evaluation Standards ใน ประเทศแอฟริกา (Africa) เยอรมัน (German) และสวิสเซอร์แลนด์ (Swiss)..	27
2.3 สังเคราะห์มาตรฐานการประเมินที่เกี่ยวข้องกับผู้ประเมิน.....	31
2.4 สรุปขั้นตอนการประเมินอภิมาน.....	48
2.5 สรุปเทคนิควิธีการประเมินอภิมานจากงานวิจัยที่เกี่ยวข้อง.....	50
2.6 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานความตรงของการประเมิน.....	53
2.7 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานอรรถประโยชน์.....	53
2.8 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานจริยธรรมในการประเมิน.....	54
2.9 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานความน่าเชื่อถือของการประเมิน.....	54
2.10 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานประสิทธิภาพของการประเมิน.....	55
2.11 สรุปงานวิจัยที่เกี่ยวกับการประเมินอภิมานงานประเมิน.....	58
2.12 ส่วนประกอบหลักในรายงานการประเมิน	60
3.1 จำนวนตัวอย่างรายงานผลการประเมินคุณภาพภายใน ระดับสถาบัน.....	75
3.2 ผลการวิเคราะห์ความตรงตามเนื้อหา (IOC) ของแบบประเมินอภิมานจาก ผู้เชี่ยวชาญ.....	81
3.3 ข้อเสนอแนะในการปรับปรุงแบบประเมินอภิมาน.....	82
3.4 ผลการวิเคราะห์ความเที่ยงของแบบประเมินอภิมานสัมประสิทธิ์แอลฟา.....	84
3.5 การประมาณค่าความแปรปรวนจากแหล่งต่างๆ ของ I X (R:P) Design ด้วย การคำนวณ G-STUDY.....	91
4.1 สังเคราะห์มาตรฐานการประเมินอภิมาน.....	104
4.2 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน มาตรฐานความตรงของการประเมิน.....	108
4.3 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานความตรงของการ ประเมิน.....	109

ตารางที่

4.4	ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน มาตรฐานอรรถประโยชน์.....	111
4.5	ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานอรรถประโยชน์.....	112
4.6	ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน มาตรฐานจริยธรรมในการประเมิน.....	113
4.7	ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานจริยธรรมในการ ประเมิน.....	114
4.8	ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน มาตรฐานความน่าเชื่อถือของนักประเมิน.....	116
4.9	ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานความน่าเชื่อถือของนัก ประเมิน.....	116
4.10	ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน มาตรฐานประสิทธิภาพของการประเมิน.....	117
4.11	ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานประสิทธิภาพของการ ประเมิน.....	118
4.12	ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนมาตรฐานในมาตรฐานการประเมิน อภิमान.....	119
4.13	ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลการประเมินอภิमान.....	120
4.14	สรุปผลการอบรมนักประเมินอภิमानทั้งสองรุ่น.....	124
4.15	ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของผลการ ประเมินอภิมานระหว่างผู้เข้ารับการอบรมนักประเมินอภิมานทั้ง 2 รุ่น.....	127
4.16	ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของผลการ ประเมินอภิมานระหว่างมหาวิทยาลัยของผู้เข้ารับการอบรมรุ่นที่ 1.....	128
4.17	ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของผลการ ประเมินอภิมานระหว่างมหาวิทยาลัยของผู้เข้ารับการอบรมรุ่นที่ 2.....	128
4.18	ผลการวิเคราะห์ความแตกต่างของค่าเฉลี่ยผลการประเมินคุณภาพรายงาน ระหว่างกลุ่มนักประเมินอภิมาน.....	129

ตารางที่	หน้า
4.19 ผลการประมาณค่าความแปรปรวนของคะแนนการประเมินอภิमानที่มีผลต่อ ค่าสัมประสิทธิ์การสรุปอ้างอิง ด้วยการประเมินรายงานผลการประเมิน คุณภาพภายใน.....	130
4.20 ผลการประมาณค่าความแปรปรวนของคะแนนการประเมินอภิमान ความคลาด เคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิงในการประเมินรายงานผลการประเมิน คุณภาพภายใน 5 แห่ง ด้วยนักประเมินอภิमानระหว่าง 2-6 กลุ่ม.....	132
4.21 ผลการประเมินอภิमानสถาบันอุดมศึกษาภาพรวม.....	134
4.22 ผลประเมินอภิमानสถาบันอุดมศึกษา จำแนกตามมาตรฐาน.....	135
4.23 ผลการประเมินอภิमानมหาวิทยาลัยของรัฐ.....	137
4.24 ผลประเมินอภิमानมหาวิทยาลัยของรัฐ จำแนกตามมาตรฐาน.....	138
4.25 ผลการประเมินอภิमानมหาวิทยาลัยเอกชน.....	140
4.26 ผลประเมินอภิमानมหาวิทยาลัยเอกชน จำแนกตามมาตรฐาน.....	141
4.27 ผลการประเมินอภิमानมหาวิทยาลัยราชภัฏ.....	143
4.28 ผลประเมินอภิमानมหาวิทยาลัยราชภัฏ จำแนกตามมาตรฐาน.....	145
4.29 ผลการประเมินอภิमानมหาวิทยาลัยเทคโนโลยีราชมงคล.....	147
4.30 ผลประเมินอภิमानมหาวิทยาลัยเทคโนโลยีราชมงคล จำแนกตามมาตรฐาน.....	148
4.31 ความถี่ ร้อยละของระดับผลการประเมินอภิमान จำแนกตามมาตรฐานและกลุ่ม สถาบัน.....	154
4.32 รายการตรวจสอบส่วนประกอบของรายงานผลการประเมินคุณภาพภายใน ของ สถาบันอุดมศึกษา ตามที่สำนักงานคณะกรรมการการอุดมศึกษากำหนด.....	155
4.33 การสังเคราะห์เนื้อหาของรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา.....	158

ศูนย์วิจัยทรัพยากร
 จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

ภาพที่	หน้า
2.1	15
2.2	21
2.3	68
2.4	69
2.5	70
3.1	74
4.1	109
4.2	111
4.3	114
4.4	116
4.5	118
4.6	120
4.7	120
4.8	130
4.9	150
4.10	150
4.11	150
4.12	150
4.13	150

บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 (สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน)[สมศ.], 2547ข) ได้กำหนดจุดมุ่งหมายและหลักการของการจัดการศึกษาที่มุ่งเน้นคุณภาพและมาตรฐาน โดยกำหนดรายละเอียดในหมวด 6 มาตรฐานและการประกันคุณภาพการศึกษา ประกอบด้วย ให้มีระบบการประกันคุณภาพการศึกษาเพื่อพัฒนาคุณภาพและมาตรฐานการศึกษาในทุกระดับ โดยมีระบบการประกันคุณภาพภายใน และระบบการประกันคุณภาพภายนอก จัดให้มีระบบ หลักเกณฑ์ และวิธีการประกันคุณภาพการศึกษา เป็นไปตามที่กฎกระทรวงกำหนดตาม มาตราที่ 47 และจัดให้มีระบบการประกันคุณภาพภายในสถานศึกษาอย่างต่อเนื่อง ให้เป็นส่วนหนึ่งของกระบวนการ การบริหารการศึกษา มีการจัดทำรายงานประจำปีเสนอต่อหน่วยงานต้นสังกัด หน่วยงานที่เกี่ยวข้อง และเปิดเผยต่อสาธารณชน เพื่อการพัฒนาคุณภาพและมาตรฐานการศึกษา และเพื่อรองรับการประกันคุณภาพภายนอก ตามมาตราที่ 48 นอกจากนี้ จัดให้มีสำนักงานรับรองมาตรฐาน และประเมินคุณภาพการศึกษา ซึ่งมีฐานะเป็นองค์การมหาชนทำหน้าที่พัฒนาเกณฑ์วิธีการประเมินคุณภาพภายนอก และทำการประเมินผลการจัดการศึกษาเพื่อให้มีการตรวจสอบคุณภาพของสถานศึกษา ในแต่ละระดับตามที่กำหนดไว้ในพระราชบัญญัตินี้ และให้มีการประเมินคุณภาพภายนอกของสถานศึกษาทุกแห่งอย่างน้อยหนึ่งครั้ง ในทุกห้าปีนับตั้งแต่การประเมินครั้งสุดท้าย และเสนอผลการประเมินต่อหน่วยงานที่เกี่ยวข้องและสาธารณชน ตามมาตราที่ 49 ทั้งนี้ข้อค้นพบ และสารสนเทศที่ได้จากการประเมินคุณภาพภายนอกนั้น มีผลกระทบโดยตรงต่อสถานศึกษาและต่อสาธารณะอีกด้วย การประเมินคุณภาพภายในเป็นกระบวนการควบคุม ตรวจสอบ ผลการดำเนินงานประกันคุณภาพ ระดับอุดมศึกษา และเตรียมรับการประเมินคุณภาพภายนอก เพื่อประเมิน และรับรองมาตรฐานคุณภาพทั้งระดับสถาบันและระดับกลุ่มสาขาวิชา จึงมีความจำเป็นต้องทำการประเมินอภิมานเพื่อวิเคราะห์ ข้อค้นพบ บรรยายและตัดสินสารสนเทศเกี่ยวกับมาตรฐานการประเมิน เพื่อช่วยให้สถานศึกษาได้เห็นจุดแข็ง จุดอ่อนของการประเมิน และตัดสินคุณค่าภายในและภายนอกได้อย่างถูกต้อง และได้รับการยอมรับจากสถานศึกษา เพื่อประกันคุณภาพการประเมินว่าได้ให้ข้อค้นพบและข้อสรุปอย่างมีประสิทธิภาพและประสิทธิผล (Stufflebeam, 2001)

การประกันคุณภาพได้เริ่มใช้เมื่อประมาณศตวรรษที่ 19 เมื่อมหาวิทยาลัยเริ่มกำหนดมาตรฐานขั้นต่ำในการรับนักศึกษา และการโอนหน่วยกิตในรายวิชาที่เท่าเทียมกันไปยังมหาวิทยาลัยอื่นๆ (Colbeck et al., 2003 อ้างถึงใน Archwamety, 2006) สหพันธรัฐของอเมริกาให้จัดทำรายชื่อมหาวิทยาลัยที่ได้รับการรับรองย้อนหลังจากปี ค.ศ.1867 และยังมีรายชื่อของโรงเรียนมัธยมปลายที่ได้รับการอนุมัติจากมหาวิทยาลัยดังกล่าวให้นักเรียนสามารถเข้าเรียนในมหาวิทยาลัยได้และหลังจากนั้นก็ยังมีหลายหน่วยงานที่ให้การรับรองมหาวิทยาลัยเพิ่มมากขึ้นตั้งแต่ปลายปี ค.ศ. 1920 ได้แก่ National Commission on Accrediting (NCA) ในปี ค.ศ. 1950 ทำหน้าที่ประสานงานกระบวนการขอการรับรองต่อมาเปลี่ยนชื่อเป็น Council on Postsecondary Accreditation (COPA) ในปี ค.ศ. 1975 ซึ่ง COPA ได้ขยายตัวมากขึ้นในปี ค.ศ. 1980 คือ มีตัวแทนอาชีพมากกว่า 70 อาชีพ และสมาคมรับรองในทั่วทั้ง 6 รัฐ (Middle State, New England, North Central, Northwest, Southern, และ Western) นอกจากนี้ในปี ค.ศ. 1993 ได้มีการก่อตั้งสมาคมชื่อ The Association of Specialized and Professional Accreditors (ASPA) ขึ้นแทนหลังจากที่ COPA เลิกดำเนินการ (Fagan and Wells, 2000 อ้างถึงใน Archwamety, 2006) ต่อมาในปี ค.ศ. 1996 ได้ก่อตั้ง The Council for Higher Education Accreditation (CHEA) (Colbeck et al., 2003 อ้างถึงใน Archwamety, 2006) เป็นหน่วยงานที่เกิดจากการรวบรวมตัวกันของมหาวิทยาลัยเอง เพื่อให้การรับรองเรื่องคุณภาพของมหาวิทยาลัย ทำหน้าที่ให้การรับรองในหลายระดับตั้งแต่ ระดับสถาบันหรือมหาวิทยาลัย ระดับวิทยาลัย/คณะ ระดับสาขาวิชา เป็นต้น และประเทศอังกฤษได้แสดงความโปร่งใสเพื่อแสดงคุณภาพและประสิทธิภาพของมหาวิทยาลัย จึงตั้ง Academic Audit Unit ทำหน้าที่ตรวจสอบกลไกประกันคุณภาพกันภายในในกลุ่มมหาวิทยาลัย นอกจากนี้รัฐบาลยังก่อตั้งหน่วยงาน Higher Education Funding Council เพื่อทำการประเมินคุณภาพ และพิจารณาจัดสรรงบประมาณ ส่วนในประเทศออสเตรเลียและนิวซีแลนด์มีความคล้ายคลึงกัน ก็คือมีการจัดตั้งหน่วยงานอิสระจากที่ประชุมอธิการบดีของทั้ง 2 ประเทศ คือ Australian Vice-Chancellors Committee (AVCC) กับ New Zealand Vice-Chancellors Committee (NZVCC) ซึ่งแยกทำงานประกันคุณภาพจากรัฐ เพื่อให้เกิดความเป็นอิสระในการดำเนินงานของสถาบันอุดมศึกษา จะเห็นได้ว่าสถาบันอุดมศึกษาจะพยายามรักษาความเป็นอิสระของตนเองไว้ โดยคิดระบบกลไกในการประกันคุณภาพเป็นของตนเอง (อมรวิรัช นาคทรพรพ, 2541)

สำหรับการประกันคุณภาพในประเทศไทยเริ่มจากในปีพ.ศ. 2540 ภายหลังจากมีการประกาศใช้รัฐธรรมนูญแห่งราชอาณาจักรไทยฉบับใหม่ และได้ประกาศใช้พระราชบัญญัติการศึกษาแห่งชาติในปี พ.ศ. 2542 แก้ไขเพิ่มเติมฉบับที่ 2 (พ.ศ. 2545) ได้กำหนดจุดมุ่งหมายและ

หลักการของการจัดการศึกษาที่มุ่งเน้นคุณภาพและมาตรฐาน โดยกำหนดรายละเอียดในหมวด 6 มาตรฐานและการประกันคุณภาพการศึกษา ประกอบด้วย “ระบบการประกันคุณภาพภายใน” และ “ระบบการประกันคุณภาพภายนอก” เพื่อเป็นกลไกในการรักษาคุณภาพและมาตรฐานของสถาบันอุดมศึกษา ดังที่กล่าวมาแล้วในข้างต้น

การประกันคุณภาพภายใน เป็นการสร้างระบบและกลไกในการควบคุม ตรวจสอบ และประเมินการดำเนินงานของสถาบันอุดมศึกษาให้เป็นที่มาของนโยบาย หลักเกณฑ์และวิธีการประกันคุณภาพของสถาบันอุดมศึกษาและหน่วยงานต้นสังกัด คือ สำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) ตามกรอบการประเมินคุณภาพภายใน 9 องค์ประกอบ 44 ตัวบ่งชี้ (สำนักงานคณะกรรมการการอุดมศึกษา, 2550) ที่พัฒนาขึ้นโดยสอดคล้องกับ มาตรฐานการอุดมศึกษา เกณฑ์มาตรฐานหลักสูตร มาตรฐานการประเมินคุณภาพภายนอกของ สมศ. และกรอบการปฏิบัติราชการตามมิติด้านต่างๆของสำนักงานคณะกรรมการพัฒนาระบบราชการ (ก.พ.ร.) ในกรณีมหาวิทยาลัยของรัฐ อีกด้วย

การประกันคุณภาพภายนอก เป็นการติดตาม ตรวจสอบ และประเมินคุณภาพ และมาตรฐานการศึกษาของสถานศึกษาซึ่งรับผิดชอบดำเนินการโดยหน่วยงานอิสระที่เรียกว่า “สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน)” หรือมีชื่อย่อว่า “สมศ.” สืบเนื่องมาจากพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542 แก้ไขเพิ่มเติม (ฉบับที่2) พ.ศ.2545 ได้กำหนดให้สถานศึกษาทุกแห่งต้องได้รับการประเมินคุณภาพภายนอกอย่างน้อย 1 ครั้ง ในทุก รอบ 5 ปี นับตั้งแต่การประเมินครั้งสุดท้าย และเสนอผลการประเมินต่อหน่วยงานที่เกี่ยวข้องและ สาธารณชนรับทราบ ซึ่ง สมศ.ได้ดำเนินการประเมินคุณภาพภายนอกครั้งแรก (พ.ศ.2544-2548) เสร็จสิ้นไปแล้ว ปัจจุบันอยู่ระหว่างประเมินคุณภาพภายนอกรอบที่สอง (พ.ศ.2549-2553) การประเมินคุณภาพภายนอกรอบที่สองมีความแตกต่างที่สำคัญจากการประเมินรอบแรก สามประการ คือ *ประการแรก* เป็นการประเมินทั้งในระดับสถาบันและกลุ่มสาขาวิชา ซึ่งสมศ.ได้ พิจารณาจัดกลุ่มสถาบันอุดมศึกษาให้สอดคล้องกันตามน้ำหนักความสำคัญ ได้ 4 กลุ่ม ดังต่อไปนี้ คือ (สมศ., 2550) 1) กลุ่มสถาบันเน้นการผลิตบัณฑิตและวิจัย 2) กลุ่มสถาบันเน้นการผลิต บัณฑิตและพัฒนาสังคม 3) กลุ่มสถาบันที่เน้นการผลิตบัณฑิตและพัฒนาศิลปะและวัฒนธรรม 4) กลุ่มสถาบันเน้นการผลิตบัณฑิต *ประการที่สอง* เป็นการประเมินตามจุดเน้นของกลุ่มสถาบัน ว่าสถาบันจะเลือกกำหนดจุดเน้นด้านใด ซึ่งมีตัวบ่งชี้เฉพาะจำนวนหนึ่งที่สถาบัน จะต้องถูก ประเมินตามจุดเน้นนอกเหนือจากตัวบ่งชี้ร่วม *ประการที่สาม* การประเมินภายนอกรอบที่สองเป็น การประเมินเพื่อรับรองมาตรฐานคุณภาพทั้งระดับสถาบันและกลุ่มสาขาวิชา นั้นหมายถึงทุกตัวบ่งชี้

จำเป็นต้องมีเกณฑ์กำกับเพื่อตัดสินคุณภาพ 3 ระดับว่า ผ่านการรับรองตามมาตรฐานคุณภาพ หรือผ่านการรับรองแบบมีเงื่อนไข หรือไม่ผ่านการรับรอง ซึ่งการประเมินในรอบสองจะมีความละเอียดและลุ่มลึกกว่าการประเมินในรอบแรก จึงจำเป็นต้องใช้ผู้ประเมินที่มีความเชี่ยวชาญเฉพาะสาขาวิชาในการตรวจประเมิน และเป็นผู้ประเมินที่มีมาตรฐานในการตรวจประเมินให้เป็นไปในทิศทางเดียวกัน (สำนักงานคณะกรรมการการอุดมศึกษา, 2550)

สำหรับแนวทางการประกันคุณภาพภายใน ระดับอุดมศึกษียึดหลักการสำคัญ 3 ประการ คือ การให้เสรีภาพทางวิชาการ (academic freedom) ความมีอิสระในการดำเนินการของสถาบัน (institutional autonomy) และความพร้อมของสถาบันที่จะรับการตรวจสอบคุณภาพจากภายนอกตามหลักการของความรับผิดชอบที่ตรวจสอบได้ (accountability) โดยสำนักงานคณะกรรมการการอุดมศึกษามีหน้าที่ส่งเสริม สนับสนุนให้สถาบันอุดมศึกษาได้จัดทำระบบการประกันคุณภาพการศึกษาภายในตามภารกิจหลักของสถาบันอุดมศึกษาให้มีประสิทธิภาพและประสิทธิผล อีกทั้งกำหนดให้มีการประเมินผล ติดตาม ตรวจสอบคุณภาพและมาตรฐานการศึกษาจากหน่วยงานต้นสังกัด ซึ่งมอบหมายให้คณะวิชาและสถานศึกษา พัฒนาระบบและกลไกการประกันคุณภาพภายใน เพื่อใช้กำกับ ตรวจสอบและประเมินคุณภาพการศึกษาให้มีประสิทธิภาพภายใต้กรอบนโยบายและหลักการที่สำนักงานคณะกรรมการการอุดมศึกษากำหนด

สถาบันอุดมศึกษาได้จัดทำรายงานการประเมินตนเอง เพื่อเผยแพร่ต่อสาธารณชนเป็นประจำทุกปี ซึ่งพบว่า ผลจากการประเมินคุณภาพรายงานการประเมินตนเองในภาพรวมอยู่ในระดับพอใช้ โดยมาตรฐานความเป็นไปได้และความเหมาะสมอยู่ในระดับพอใช้ มาตรฐานด้านความถูกต้องอยู่ในระดับควรปรับปรุง ส่วนด้านอรรถประโยชน์อยู่ในระดับดี จากงานวิจัยของ วารุณี วัฒนโชคดี (2550) นอกจากนี้สำนักงานคณะกรรมการการอุดมศึกษาได้มีการจัดอบรมเพื่อพัฒนา ผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษา ให้มีความรู้ และทักษะในการประเมินคุณภาพภายใน และเพื่อเป็นการพัฒนาระบบและกลไกในการประกันคุณภาพภายในให้มีประสิทธิภาพมากยิ่งขึ้น จึงได้กำหนดให้มีการประเมินคุณภาพภายใน ระดับสถาบัน โดยผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษาที่ผ่านการอบรมและเป็นบุคคลภายนอกสถาบันสัดส่วน 2 ใน 3 เพื่อให้เกิดความโปร่งใสในการประเมิน อีกทั้งได้รับมุมมองที่หลากหลายมากขึ้น

การควบคุมคุณภาพ การประเมิน เพื่อเป็นการสะท้อนให้เห็นถึงคุณภาพของการประเมินคุณภาพภายใน ระดับอุดมศึกษา จึงจำเป็นต้องมีการตรวจสอบงานการประเมิน หรือทำการประเมินซ้ำ ด้วยวิธีวิทยาการประเมิน ในศาสตร์ทางการประเมินมีความก้าวหน้าด้วยการทำการประเมินอภิมาน ซึ่ง Stufflebeam (2001) เสนอว่า นักประเมินสามารถทำการประเมินอภิมาน

อย่างเป็นระบบได้ เนื่องจากการประเมินอภิมานเป็นกระบวนการวิเคราะห์ การค้นพบ และการประยุกต์เพื่อบรรยายสารสนเทศ และการตัดสินสารสนเทศเกี่ยวกับมาตรฐานการประเมินใน ด้าน อรรถประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง และธรรมชาติระบบ สมรรถนะ ความซื่อสัตย์ การยอมรับ และความรับผิดชอบต่อสังคม เพื่อเป็นแนวทางในการประเมินและ รายงานจุดแข็ง จุดอ่อนสู่สาธารณชน ซึ่งการประเมินอภิมานความก้าวหน้า (formative metaevaluations) เป็นการยอมรับและการทำการประเมิน เพื่อช่วยนักประเมินวางแผน ดำเนินการ ปรับปรุง แปลความและรายงานการศึกษาประเมิน ส่วนการประเมินอภิมานสรุปรวม (summative metaevaluations) ดำเนินการติดตามการประเมินเพื่อช่วยผู้รับผลการประเมินเห็น จุดแข็งและจุดอ่อนของการประเมิน และตัดสินคุณค่าภายในและภายนอก การประเมินอภิมาน เป็นที่เปิดเผย ได้รับการยอมรับและความสนใจของสถาบัน เพื่อประกันว่าการประเมินให้ข้อค้นพบ และข้อสรุป ซึ่งการทำการประเมินอย่างต่อเนื่องเพื่อปรับปรุง และการบริหารสถาบัน ให้ระบบการ ประเมินอย่างมีประสิทธิภาพและประสิทธิผล นักประเมินมืออาชีพมีหน้าที่ทำการประเมินอภิมาน อย่างเข้มงวดมากยิ่งขึ้น แต่จำเป็นต้องเพิ่มเครื่องมือและกระบวนการที่จะประยุกต์ใช้มาตรฐาน และหลักการของแนวปฏิบัติการประเมินที่ดี

Scriven (2003) และStufflebeam (2002) ได้สรุปวัตถุประสงค์ของการประเมิน อภิมาน คือ เพื่อตรวจสอบ ควบคุม และรับรองคุณภาพของการประเมินในกรณีหลายโครงการหรือ ผู้ประเมินหลายคน รวมถึงเพื่อสังเคราะห์สรุปผลการประเมิน สำหรับวิธีที่ใช้ในการประเมินอภิมาน Davidsion (2005) ได้เสนอทางเลือกในการทำการประเมินอภิมานไว้ 3 ทางเลือก ดังนี้ **ทางเลือก แรก** เป็นการประเมินตามมาตรฐานการประเมินโครงการ ของ Joint Committee on Standards for Educational Evaluation (1994) โดยทำแบบตรวจสอบรายการประเมินอภิมาน (meta evaluation checklist) มีประโยชน์ในการใช้ตรวจสอบรายการตามจุดเน้น 4 ด้าน คือ ด้านความ เหมาะสม (propriety) ความเป็นไปได้ (feasibility) อรรถประโยชน์ (utility) และด้านความถูกต้อง (accuracy) **ทางเลือกที่สอง** เป็นวิธีการประเมินโดยใช้ทีมประเมินที่แตกต่างกัน ซึ่งมีความเป็น อิสระของการประเมิน เป็นการทำการประเมินในสิ่งเดียวกัน และมีการรายงานเพื่อเปรียบเทียบผล การประเมิน พิจารณาส่งที่เหมือนกันและหากพบสิ่งแตกต่างกัน จะค้นหาข้อสรุปเดียวกันหรือข้อ ค้นพบที่ถูกต้อง **ทางเลือกสุดท้าย** เป็นวิธีการผสม โดยวิธีนี้นักประเมินอภิมานจะทำการประเมิน และมีการตรวจสอบไขว้ผลจากข้อค้นพบที่น่าสงสัย

การประเมินอภิมาน ด้วยแบบตรวจสอบสำหรับการประเมินอภิมานรายงานประเมิน โครงการ (program evaluations meta-evaluation checklist) ที่พัฒนาขึ้นโดย Stufflebeam (1999)

โดยอิงเกณฑ์มาตรฐานการประเมินโครงการของ The Joint Committee on Standards for Educational Evaluation พบว่า รายงานการประเมินคุณภาพภายนอกส่วนใหญ่มีคุณภาพตามมาตรฐานการประเมินอภิมาน โดยมาตรฐานด้านอรรถประโยชน์ ที่ประกันได้ว่าการประเมินคุณภาพภายนอกมีสารสนเทศสมบูรณ์ ชัดเจน มีคุณค่า ทันเวลา และเป็นประโยชน์ต่อสถานศึกษา ในระดับที่ดีถึงดีมาก มาตรฐานด้านความเป็นไปได้ ที่มีการปฏิบัติที่เหมาะสม ใช้เวลาและทรัพยากรประหยัด และคำนึงถึงอำนาจการเมืองในสถานศึกษาในระดับพอใช้ มาตรฐานด้านความเหมาะสมชอบธรรม ที่สอดคล้องกับจรรยาบรรณและสิทธิมนุษยชน และมาตรฐานด้านความถูกต้อง อยู่ในระดับพอใช้ ซึ่งผลการประเมินอภิมานสอดคล้องกับงานวิจัยของวารุณี ลัภนโชคดี (2550) พบว่า ผลการประเมินอภิมานรายงานการประเมินคุณภาพภายนอกรอบแรก ของสถาบันอุดมศึกษาทุกประเภทมีคุณภาพในภาพรวมใกล้เคียงกันคือ อยู่ในระดับพอใช้ โดยมาตรฐานด้านอรรถประโยชน์อยู่ในระดับดี มาตรฐานด้านความเป็นไปได้ มาตรฐานด้านความเหมาะสมชอบธรรม และมาตรฐานด้านความถูกต้อง อยู่ในระดับพอใช้ และงานวิจัยของนางลักษณ วัชรชัย และวรรณี เจตจำนงนุช (2548) ใช้การประเมินอภิมานเป็นการประเมินอภิมานสรุปรวม (summative meta-evaluation) จากรายงานการประเมินคุณภาพภายนอก ระดับอุดมศึกษา ซึ่งมาตรฐานในการประเมินอภิมาน ประกอบด้วย มาตรฐานย่อย 30 มาตรฐาน และมาตรฐานย่อยบางมาตรฐานนั้น สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) หรือสมศ. มิได้กำหนดให้คณะผู้ประเมินภายนอกต้องปฏิบัติและให้แสดงไว้ในรายงานการประเมิน ได้แก่ มาตรฐานย่อย F2 ความอยู่รอดทางการเมือง และ P8 ความรับผิดชอบด้านการเงิน น่าจะเป็นมาตรฐานที่ไม่ค่อยเกี่ยวข้องกับการประเมินคุณภาพภายนอก เนื่องจากเป็นการประเมินตามข้อกำหนดในพระราชบัญญัติการศึกษาแห่งชาติ ที่สถาบันทุกสถาบันรับรู้และยอมให้มีการปฏิบัติอยู่แล้ว ซึ่งมีผลให้การประเมินในมาตรฐานย่อยดังกล่าวไม่ดีและส่งผลกระทบต่อผลการประเมินในมาตรฐานหลักอีกด้วย (นางลักษณ วัชรชัย และวรรณี เจตจำนงนุช, 2548; วารุณี ลัภนโชคดี, 2550)

แนวทางการพัฒนางานประเมินให้มีความน่าเชื่อถือและเป็นที่ยอมรับโดยสากล จึงควรพิจารณาบททวน เพื่อพัฒนามาตรฐานในการประเมินอภิมานที่เหมาะสมกับสภาพและบริบทการประเมินในประเทศไทย โดยการสังเคราะห์จากมาตรฐานการประเมิน Program Evaluation Standards ตามแนวทางของ the Joint Committee on Standard for Educational Evaluation (1994) ประกอบด้วย 4 มาตรฐานหลัก 30 มาตรฐานย่อย The African Evaluation guidelines 2002 (AEG) ที่พัฒนาขึ้นโดยสมาคมการประเมินของประเทศแอฟริกา (African Evaluation Association) นำโดย Patel (2002) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 30 มาตรฐานย่อย มาตรฐานการ

ประเมินในบริบทของประเทศสวิตเซอร์แลนด์ เรียกว่า SEVAL's Evaluation Standards พัฒนาโดยสมาคมการประเมินของประเทศสวิตเซอร์แลนด์ (Swiss Evaluation Society) นำโดย Widmer จากมหาวิทยาลัยซูริค (Zurich University) กำหนดมาตรฐานการประเมินหลัก 4 มาตรฐาน และ 28 มาตรฐานย่อย (Laubli Loud, 2004) The DeGEval's Standards (2003) ที่พัฒนาโดยสมาคมการประเมินของประเทศเยอรมัน (German Evaluation Society) นำโดย Beywl (2003) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 25 มาตรฐานย่อย หลักการและแนวทางการประเมินสำหรับผู้ประเมิน Guiding Principles for Evaluators ของสมาคมประเมินแห่งสหรัฐอเมริกา (American Evaluation Association [AEA], 2003) ทักษะที่จำเป็นในการประเมิน Essential Skills Series in Evaluation ของสมาคมประเมินแห่งแคนาดา (Canadian Evaluation Society, 1999) และสมรรถนะที่จำเป็นสำหรับผู้ประเมินโครงการ Essential Competencies for Program Evaluators (Ghere & et al., 2006) โดยยึดกรอบมาตรฐานการประเมิน 5 มาตรฐาน ของ Davidson (2005) ได้แก่ 1) ความตรงของการประเมิน (validity) 2) รรถประโยชน์ (utility) 3) จริยธรรมในการประเมิน (ethicality) 4) ความน่าเชื่อถือของผู้ประเมิน (credibility) และ 5) ประสิทธิภาพต้นทุน (cost-effectiveness) นอกจากนี้มาตรฐานการประเมินอภิมาน 5 มาตรฐาน ดังกล่าวได้นำไปใช้เป็นกรอบในการประเมินทุนวิจัยของรัฐบาล เปรียบเทียบระหว่างประเทศใน 16 ประเทศ เพื่อใช้พิจารณาความเหมาะสมของโมเดลการประเมินอีกด้วย (Coryn, Hattie, Scriven, & Hartmann, 2007; Coryn & Scriven, 2007)

สำหรับการวิจัยในครั้งนี้จะทำการพัฒนามาตรฐานการประเมินอภิมาน 5 มาตรฐาน ตามแนวคิดของ Davidson (2005) เนื่องจากเป็นมาตรฐานที่สามารถประเมินได้ครอบคลุมมาตรฐานการประเมิน Program Evaluation Standards ของ Stufflebeam และการประเมินงานประเมินจากเอกสารรายงาน ต้องทำอย่างถูกต้องและสรุปอย่างสมเหตุสมผล มีประโยชน์ต่อผู้รับการประเมินและผู้ที่เกี่ยวข้อง ทำการประเมินอย่างมีจริยธรรม ตามกฎระเบียบ เชี่ยวชาญและเหมาะสม น่าเชื่อถือต่อผู้ที่เกี่ยวข้อง รวมถึงประหยัด รวดเร็ว และไม่ยุ่งยาก เพื่อให้เหมาะสมกับบริบทประเทศไทย ที่ได้จากการสังเคราะห์จากการประยุกต์ใช้มาตรฐานการประเมินในบริบทของประเทศต่างๆ ใน 7 แนวทางการประเมิน โดยดำเนินการพัฒนามาตรฐาน ตัวบ่งชี้ และเกณฑ์ในการประเมินอภิมานที่เหมาะสม เพื่อพัฒนาจัดทำแบบประเมินอภิมาน และดำเนินการพัฒนาหลักสูตรการฝึกอบรมนักประเมินอภิมานเพื่อให้มีความรู้ ความสามารถในการประเมินอภิมานคุณภาพรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา ตามกรอบมาตรฐานการประเมินอภิมานที่พัฒนาขึ้น สำหรับใช้เป็นแนวทางในการพัฒนาระบบการประเมินคุณภาพภายในและส่งเสริมให้มีการประเมินอภิมานในสถาบันอุดมศึกษาต่อไป

คำถามการวิจัย

1. มาตรฐานการประเมินอภิमान (Meta-evaluation Standards) สำหรับประเมินอภิमान ระดับอุดมศึกษาในประเทศไทย ควรมีลักษณะเป็นอย่างไร
2. มาตรฐานการประเมินมีคุณภาพด้านความตรงและความเที่ยงมากน้อยเพียงใด
3. การประยุกต์ใช้มาตรฐานการประเมินอภิमानในการประเมินรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา จะมีผลการประเมินคุณภาพสถาบันอุดมศึกษาเป็นอย่างไร

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนามาตรฐานการประเมินอภิमान สำหรับประเมินรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาในประเทศไทย
2. เพื่อตรวจสอบคุณภาพของมาตรฐานการประเมินอภิमान ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา
3. เพื่อประยุกต์ใช้มาตรฐานการประเมินอภิमानในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา ระดับอุดมศึกษา

ขอบเขตของการวิจัย

การวิจัยในครั้งนี้ เป็นการพัฒนามาตรฐานการประเมินอภิमानโดยใช้หลักแนวคิดกรอบ 5 มาตรฐานของ Davidson (2005) ซึ่งสังเคราะห์ด้ว้งซึ่งการประเมินจาก 7 แนวทางในการประเมิน คือ 1) Program Evaluation Standards (1994) 2) The African Evaluation guidelines 2002 (AEG) 3) SEVAL's Evaluation Standards (2004) 4) The DeGEval's Standards (2003) 5) Guiding Principles for Evaluators (2003) 6) Essential Skills Series in Evaluation (1999) และ 7) Essential Competencies for Program Evaluators (2006) เพื่อใช้ในการประเมินรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา โดยผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับมาตรฐานการประเมิน การประเมินอภิमान ขั้นตอนและกระบวนการการประเมินอภิमान ซึ่งได้พัฒนามาตรฐานการประเมินอภิमानและประยุกต์ใช้มาตรฐานการประเมิน

การตรวจสอบมาตรฐานการประเมินอภิमान ด้วยการพัฒนาเครื่องมือในการประเมินอภิमान คู่มือการประเมินอภิमान หลักสูตรการอบรมนักประเมินอภิमान โดยให้ผู้เชี่ยวชาญ 5 ท่าน พิจารณาคุณภาพ และทดลองใช้เครื่องมือโดยนำไปใช้ในการอบรมนักประเมินอภิमान 2 รุ่น จำนวน 28 คน เพื่อใช้ประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา ระดับอุดมศึกษา

ซึ่งเป็น การประเมินอภิมานแบบสรุปรวม (summative meta-evaluation) และหาคุณภาพด้านความเที่ยงด้วยสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) ระหว่างนักประเมินอภิมานต่างกัน ที่ประเมินรายงานฯฉบับเดียวกัน และหาความตรงตามโครงสร้าง (construct validity) ด้วยเทคนิคกลุ่มรู้จัก (known group technique) เทียบจากผู้ทรงคุณวุฒิด้านการประกันคุณภาพจำนวน 10 ท่าน กับผลการประเมินอภิมานจากนักประเมินอภิมานที่ผ่านการอบรมแล้ว โดยใช้แหล่งข้อมูลหลักเป็นรายงานผลการประเมินคุณภาพภายในและรายงานการประเมินตนเอง (SAR) ของสถาบันอุดมศึกษา ที่เป็นไปตามเงื่อนไขที่กำหนด คือ ได้รับความยินยอมให้ข้อมูลการประเมินจากหน่วยงานต้นสังกัดและสถาบัน และได้รับการประเมินคุณภาพภายใน ระดับสถาบันประจำปีการศึกษา 2550 โดยมีการเผยแพร่รายงานการประเมินครบถ้วน สมบูรณ์ ทั้ง รายงานผลการประเมินคุณภาพภายใน และ SAR จำนวนรวม 50 ฉบับ จำแนกเป็น มหาวิทยาลัยของรัฐ 8 ฉบับ มหาวิทยาลัยเอกชน 19 ฉบับ มหาวิทยาลัยราชภัฏ 18 ฉบับ และมหาวิทยาลัยเทคโนโลยีราชมงคล 5 ฉบับ และผลจากการประยุกต์ใช้มาตรฐานการประเมินอภิมาน นำมาวิเคราะห์หาคุณภาพของรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ตามมาตรฐานการประเมินอภิมาน

คำจำกัดความที่ใช้ในการวิจัย

มาตรฐาน หมายถึง ข้อกำหนดเกี่ยวกับคุณลักษณะหรือระดับคุณภาพที่พึงประสงค์ เพื่อใช้เป็นหลักเทียบเคียง สำหรับส่งเสริม กำกับดูแล ตรวจสอบ ประเมินผล และประกันคุณภาพ

การประเมินอภิมาน หมายถึง การตัดสินคุณค่าของรายงานผลการประเมินเทียบกับเกณฑ์ มาตรฐานการประเมินอภิมาน 5 มาตรฐาน ที่พัฒนาขึ้น โดยทำการประเมินจากรายงานผลการประเมินคุณภาพภายใน ที่ผ่านการประเมินจากผู้ประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาของสำนักงานคณะกรรมการการอุดมศึกษาแล้ว เพื่อวิเคราะห์หาจุดแข็ง จุดอ่อนของรายงานผลประเมินและเพื่อควบคุมคุณภาพของการประเมินคุณภาพภายในระดับอุดมศึกษา

การพัฒนามาตรฐานการประเมินอภิมาน หมายถึง การพัฒนาข้อกำหนดหรือคุณลักษณะ ที่พึงประสงค์ของงานประเมิน ที่ใช้ในการควบคุมคุณภาพการประเมินตามมาตรฐานการประเมินอภิมานที่ทำให้สารสนเทศที่ได้จากการประเมินเกิดความยุติธรรม และครอบคลุมมาตรฐานหลัก 5 มาตรฐาน คือ มาตรฐานความตรงของการประเมิน (validity) มาตรฐานอรรถประโยชน์ (utility) มาตรฐานจริยธรรมในการประเมิน (ethicality) มาตรฐานความน่าเชื่อถือของนักประเมิน (credibility) และมาตรฐานประสิทธิผลต้นทุน (cost-effectiveness) ที่สังเคราะห์ขึ้นจากมาตรฐานการประเมินที่

เกี่ยวข้องใน 7 แนวทางการประเมิน คือ PES (1994), AEG (2002), SEVAL (2004), DeGEval (2003), AEA (2003), CES (1999), และ ECPE (2006) ประกอบด้วย 38 ตัวบ่งชี้ โดยการพัฒนาคุณภาพมาตรฐานการประเมินอภิมานจากการพิจารณาของผู้เชี่ยวชาญด้านการประกันคุณภาพ เพื่อปรับปรุงให้เหมาะสมกับบริบทการประเมินในประเทศไทย

มาตรฐานความตรงของการประเมิน หมายถึง คุณลักษณะของงานประเมิน ที่มีการจัดการเอกสารในการประเมิน การวิเคราะห์บริบท กำหนดวัตถุประสงค์และออกแบบการประเมิน การเก็บรวบรวมสารสนเทศ วิเคราะห์ แปลผลและสรุปสารสนเทศอย่างมีคุณภาพ และตรวจสอบการประเมินอย่างถูกต้องสมเหตุสมผล ประกอบด้วย 13 ตัวบ่งชี้

มาตรฐานอรรถประโยชน์ หมายถึง คุณลักษณะของงานประเมิน ที่เป็นประโยชน์ต่อผู้มีส่วนได้ส่วนเสีย และผู้ที่เกี่ยวข้อง มีการตัดสินใจคุณค่า และรายงานผลการประเมินอย่างชัดเจน มีการเผยแพร่ผลการประเมินทันความต้องการของผู้ใช้ผลการประเมิน และเป็นประโยชน์ในการวางแผน ปรับปรุงการดำเนินงานในทิศทางที่เสนอแนะ ประกอบด้วย 10 ตัวบ่งชี้

มาตรฐานจริยธรรมในการประเมิน หมายถึง คุณลักษณะของงานประเมินที่มีการกำหนดวิธีการประเมินไว้เหมาะสมกับสภาพจริง ได้รับการยอมรับ มีการคำนึงถึงผู้ที่เกี่ยวข้อง หลายกลุ่ม มีการปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงสิทธิมนุษยชน และประโยชน์สาธารณะ ทำการประเมินที่มีความถูกต้องยุติธรรมต่อผู้รับผลการประเมิน รวมถึงเปิดเผยผลการประเมินแก่ผู้ที่ได้รับผลกระทบจากการประเมิน ประกอบด้วย 9 ตัวบ่งชี้

มาตรฐานความน่าเชื่อถือของนักประเมิน หมายถึง คุณลักษณะของงานประเมินที่ได้มาจากนักประเมินที่มีสมรรถนะในการประเมิน ได้รับการยอมรับ และไม่มีผลประโยชน์ทับซ้อนเพื่อสร้างความน่าเชื่อถือของข้อค้นพบและสารสนเทศที่ได้รับจากการประเมิน ประกอบด้วย 4 ตัวบ่งชี้

มาตรฐานประสิทธิผลต้นทุน หมายถึง คุณลักษณะของงานประเมิน ที่คำนึงถึงความคุ้มค่าของทรัพยากรที่จำเป็นในการประเมิน มีค่าใช้จ่ายในการประเมินอย่างสมเหตุสมผล มีหลักฐานชัดเจนและมีความโปร่งใสตรวจสอบได้ของต้นทุนการประเมิน ประกอบด้วย 2 ตัวบ่งชี้

แบบประเมินอภิมาน หมายถึง เครื่องมือที่ผู้วิจัยสร้างขึ้น เพื่อใช้พิจารณาตรวจสอบตัดสินคุณค่า และประเมินค่า รายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาที่เป็นไปตามกรอบมาตรฐานการประเมินอภิมาน 5 มาตรฐาน แบ่งออกเป็น 3 ตอน คือ ตอนที่ 1 ข้อมูลทั่วไปของสถานศึกษา ตอนที่ 2 แบบประเมินเป็นชนิดกำหนดเกณฑ์การให้คะแนน (rubric scoring) 4 ระดับ (1-4) จำนวนทั้งหมด 38 ข้อ จำแนกตามมาตรฐาน และตอนที่ 3 ข้อเสนอแนะในการนำผลการประเมินไปใช้ประโยชน์

การประยุกต์ใช้มาตรฐานการประเมินอภิमान หมายถึง การนำมาตรฐานการประเมินอภิमानที่พัฒนาขึ้น ไปพัฒนาเครื่องมือในการประเมินอภิमान คู่มือการประเมินอภิमान และหลักสูตรในการฝึกอบรม เพื่อพัฒนานักประเมินอภิमानที่สามารถประเมินอภิमानรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา

รายงานผลการประเมินคุณภาพภายใน หมายถึง รายงานผลการประเมินคุณภาพการจัดการศึกษา การติดตาม การตรวจสอบคุณภาพและมาตรฐานการศึกษาของสถาบันอุดมศึกษา ที่ประเมินโดยผู้ประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา ที่ผ่านการอบรมของสำนักงานคณะกรรมการการอุดมศึกษา ตามมาตรฐานและเกณฑ์การประเมินที่สำนักงานคณะกรรมการการอุดมศึกษากำหนด

ผลการประเมินคุณภาพภายใน หมายถึง ระดับของผลการประเมินคุณภาพภายใน ระดับสถาบัน ที่ได้จากการประเมินคุณภาพภายในเทียบกับเกณฑ์และมาตรฐานการประเมินของสำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) ซึ่งดำเนินการประเมินโดยทีมผู้ประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาที่ผ่านการอบรมโครงการฝึกอบรมผู้ประเมินคุณภาพภายในระดับอุดมศึกษา และมีคณะกรรมการประเมินคุณภาพภายในตามหลักเกณฑ์ที่ สกอ. กำหนดไว้

สถาบันอุดมศึกษา หมายถึง วิทยาลัย สถาบัน มหาวิทยาลัย หรือ หน่วยงานที่เรียกชื่ออย่างอื่น ทั้งของรัฐและเอกชนที่จัดการศึกษาระดับอุดมศึกษาต่ำกว่าปริญญาและ/หรือ ระดับปริญญา ได้แก่ สถานศึกษาระดับอุดมศึกษา ในสังกัดสำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ และสถาบันเฉพาะทางสังกัดกระทรวงอื่นๆ

การตรวจสอบคุณภาพเครื่องมือประเมินอภิमान หมายถึง เครื่องมือในการประเมินอภิमानที่ผ่านการตรวจสอบคุณภาพ ด้านความตรงตามเนื้อหา (content validity) จากผู้เชี่ยวชาญ จำนวน 5 ท่าน ความตรงเชิงโครงสร้าง (construct validity) ด้วยเทคนิคกลุ่มรู้จัก (known group technique) จากผลการประเมินของนักประเมินอภิमानเทียบกับคุณภาพของรายงานที่บ่งชี้โดยผู้ทรงคุณวุฒิด้านการประกันคุณภาพ และความเที่ยงแบบสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient)

ประโยชน์ที่ได้รับ

ผลจากการวิจัยในครั้งนี้ทำให้เกิดประโยชน์ต่อการพัฒนาระบบการประเมินคุณภาพภายใน ของสถานศึกษา ในระดับอุดมศึกษา ดังต่อไปนี้

1. ได้มาตรฐานการประเมินอภิमान 5 มาตรฐาน ที่สังเคราะห์ที่ตัวบ่งชี้จากมาตรฐานการประเมินใน 7 แนวทางการประเมินที่สอดคล้องกับ พระราชบัญญัติการศึกษาแห่งชาติ มาตรฐาน

การศึกษาชาติ และประยุกต์ใช้ให้เหมาะสมกับการประเมินในบริบทของประเทศไทยและสามารถนำไปปรับใช้ในการประเมินด้านต่างๆ

2. ได้ผลการตัดสินคุณภาพของรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา ที่ได้รับการประเมินโดยผู้ประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา ของสกอ. พิจารณาเทียบกับมาตรฐานและเกณฑ์การประเมินอภิมานที่พัฒนาขึ้น

3. มาตรฐานที่ได้สามารถดูแล กำกับ ติดตามกระบวนการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา และการเขียนรายงานผลการประเมินคุณภาพภายใน ให้มีประสิทธิภาพมากขึ้น

4. สถานศึกษาระดับอุดมศึกษา สามารถเห็นจุดแข็ง จุดอ่อนของการประเมินคุณภาพภายใน และสามารถตัดสินคุณค่าภายในและภายนอกของการประเมินได้อย่างถูกต้อง รวมทั้งได้รับการยอมรับผลการประเมินคุณภาพภายใน จากสถานศึกษาและสาธารณชน

5. หลักสูตรฝึกอบรมสามารถประกันได้ว่าข้อค้นพบที่ได้จากการประเมินคุณภาพภายใน ระดับอุดมศึกษา โดยผู้ประเมินคุณภาพภายใน ที่ผ่านการอบรมจากสำนักงานคณะกรรมการการอุดมศึกษา มีประสิทธิภาพและประสิทธิผล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยเสนอแนวคิด ทฤษฎี เอกสารงานวิจัยที่เกี่ยวข้องกับการวิจัยในครั้งนี้ โดยแบ่งเป็น 5 ตอน คือ ตอนที่ 1 แนวคิดทฤษฎีเกี่ยวกับการประกันคุณภาพและงานวิจัยที่เกี่ยวข้อง ตอนที่ 2 แนวคิดเกี่ยวกับมาตรฐานการประเมิน เกณฑ์การประเมินและงานวิจัยที่เกี่ยวข้อง ตอนที่ 3 แนวคิดทฤษฎีเกี่ยวกับการประเมินอภิมานและงานวิจัยที่เกี่ยวข้อง ตอนที่ 4 แนวทางการเขียนรายงานการประเมินและงานวิจัยที่เกี่ยวข้อง และตอนที่ 5 กรอบแนวคิดในการวิจัย ดังรายละเอียดดังต่อไปนี้

ตอนที่ 1 แนวคิดทฤษฎีเกี่ยวกับการประกันคุณภาพ

แนวคิดทฤษฎีเกี่ยวกับการประกันคุณภาพในส่วนนี้แบ่งออกเป็น 3 หัวข้อ ได้แก่ 1) การประกันคุณภาพการศึกษา 2) ความสัมพันธ์ระหว่างมาตรฐานการศึกษาของชาติกับมาตรฐานการประเมินคุณภาพการศึกษา และ 3) งานวิจัยที่เกี่ยวข้อง ดังมีรายละเอียดต่อไปนี้

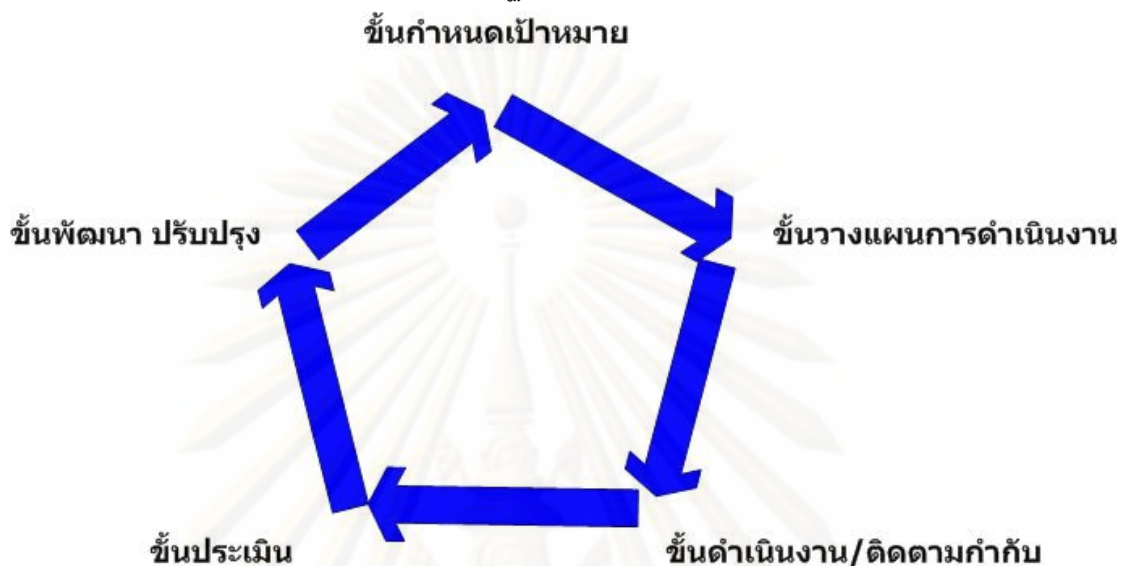
1. การประกันคุณภาพการศึกษา

การประกันคุณภาพ มีนักวิชาการได้ให้คำนิยามไว้หลายประการ ได้แก่ ‘การประกันคุณภาพ’ มาจากคำว่า ‘การประกัน’ และคำว่า ‘คุณภาพ’ คำว่า ‘การประกัน’ หมายถึง กระบวนการที่รับรองด้วยความเชื่อมั่นในเรื่องใดเรื่องหนึ่ง ส่วนคำว่า ‘คุณภาพ’ นั้นนักวิชาการให้นิยามไว้แตกต่างกันหลายแบบ ทบวงมหาวิทยาลัย (2539) สรุปนิยามของคำว่า ‘คุณภาพ’ ได้เป็นสั้นๆ นัยแรก คุณภาพ คือ การมีมาตรฐานความเป็นเลิศ โดยที่ต้องมีการกำหนดมาตรฐานไว้ก่อน ทั้งนี้มาตรฐานอาจมีการเปลี่ยนแปลงได้ตามความเหมาะสมกับสภาพทางสังคม นัยที่สอง คุณภาพ คือ ความตรงตามวัตถุประสงค์ที่กำหนดไว้ นัยที่สาม คุณภาพ คือ ประสิทธิภาพการดำเนินงานที่มีการใช้ทรัพยากรอย่างคุ้มค่า และนัยที่สี่ คุณภาพ คือ การพัฒนาให้ดียิ่งขึ้น นอกจากนี้ สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) (2547) ให้นิยามว่า คุณภาพการศึกษา คือ คุณลักษณะของผู้เรียนที่มีผลสัมฤทธิ์ทางการเรียนสูงและมีคุณภาพดี คุณภาพของผู้เรียนดังกล่าวขึ้นอยู่กับความเอาใจใส่ของพ่อแม่ ผู้ปกครองและชุมชน ระบบการบริหารที่ได้มาตรฐานและระบบการเรียนการสอนที่มีคุณภาพ อีกทั้งเชื่อว่า ความเอาใจใส่ของพ่อแม่ ผู้ปกครองและชุมชนน่าจะส่งผลโดยอ้อมต่อระบบการบริหาร และระบบการเรียนการสอนที่ได้มาตรฐาน และส่งผลโดยตรงต่อการได้มาซึ่งคณะกรรมการสถานศึกษาที่เชี่ยวชาญ คณะกรรมการ

สถานศึกษาที่เชี่ยวชาญน่าจะได้ส่งผลโดยตรงต่อการมีผู้บริหารและครูมืออาชีพ และน่าจะได้ส่งผลโดยอ้อมต่อมาตรฐานของระบบบริหารและการจัดการเรียนการสอน อีกทั้งเชื่อว่า ผู้บริหารและครูมืออาชีพ จะนำไปสู่ระบบการบริหารและระบบการเรียนการสอนที่มีคุณภาพและได้มาตรฐานอันจะนำไปสู่การพัฒนาผู้เรียนให้มีผลสัมฤทธิ์ของการเรียนสูงและมีคุณภาพดี เป็นคนดี มีความสามารถ และมีความสุข สนุกกับการเรียนรู้ตลอดชีวิต

การศึกษาได้นำแนวคิดเรื่องการประกันคุณภาพมาใช้ในช่วงคริสต์ศตวรรษ 1990s นักการศึกษากำหนดว่าเป้าหมายสำคัญในการจัดการศึกษาและการจัดการเรียนการสอน คือคุณภาพ และให้นิยามคำว่าคุณภาพของโรงเรียนโดยเชื่อมโยงกับเป้าหมายของการจัดการศึกษาว่า สถานศึกษาที่มีคุณภาพ หมายถึง สถานศึกษาที่กำหนดจุดมุ่งหมาย/เป้าหมายของสถานศึกษาอย่างชัดเจน และดำเนินงานอย่างมีประสิทธิภาพ มีการประเมินผลการดำเนินงานว่าได้ผลตามเป้าหมาย/มาตรฐาน และมีการใช้ผลการประเมินเพื่อพัฒนากระบวนการดำเนินงานให้มีประสิทธิภาพ เกิดประสิทธิผล สามารถบรรลุจุดมุ่งหมายที่กำหนดไว้ การวัดคุณภาพของสถานศึกษาจึงต้องวัดหลายมิติทั้งในมิติของเป้าหมาย กระบวนการดำเนินงาน และผลผลิต สำหรับกระบวนการดำเนินงานอย่างมีระบบดังกล่าวข้างต้น เพื่อกำกับติดตามให้แน่ใจว่าสถานศึกษามีคุณภาพ เรียกว่าการประกันคุณภาพการศึกษา (Segers and Dochy, 1996; Webstock, 1997) ซึ่งกระบวนการประกันคุณภาพการศึกษามีจุดมุ่งหมายสำคัญ คือการพัฒนาคุณภาพการศึกษาให้ดีขึ้น เพื่อให้ผู้ปกครอง และผู้เกี่ยวข้องมั่นใจได้ว่าสถานศึกษาจัดการศึกษาได้อย่างมีประสิทธิภาพ การประกันคุณภาพเป็นเครื่องมือสำคัญอย่างหนึ่งที่ช่วยให้ผู้บริหารทราบว่าการดำเนินงานของสถานศึกษามีความโปร่งใส และตรวจสอบได้ ช่วยให้สถานศึกษาได้ทราบจุดแข็ง จุดอ่อนและสามารถกำหนดแนวทางการพัฒนาสถานศึกษาได้อย่างมีประสิทธิภาพ งานสำคัญของกระบวนการประกันคุณภาพการศึกษา คือ การประเมินภายนอก และการประกันคุณภาพภายในสถานศึกษา กล่าวคือ ในการจัดการศึกษาให้มีคุณภาพนอกจากสถานศึกษาต้องจัดให้มีระบบประกันคุณภาพภายในสถานศึกษาแล้ว ยังต้องจัดให้มีระบบการประเมินตรวจสอบโดยหน่วยงานภายนอกที่เป็นกลางทำหน้าที่ประเมินคุณภาพสถานศึกษา เพื่อกระตุ้นให้การดำเนินการประกันคุณภาพภายใน เพื่อพัฒนาสถานศึกษา เป็นไปอย่างต่อเนื่องและสม่ำเสมอ ทั้งนี้การประเมินภายนอกยังดำเนินการได้เป็น 2 แบบ คือ การประเมินโดยหน่วยงานต้นสังกัด/หน่วยงานสนับสนุนงบประมาณดำเนินงาน และการประเมินโดยหน่วยงานภายนอก Segers และ Dochy (1996) ให้หลักการว่ากระบวนการประกันคุณภาพภายในสถานศึกษาทำได้ทุกระดับทั้งในระดับสถานศึกษา

ระดับหลักสูตร ระดับรายวิชา การดำเนินงานอาจมีขั้นตอนแตกต่างกันได้ แต่ขั้นตอนสำคัญมี 5 ขั้นตอนที่ต้องดำเนินการต่อเนื่องกันไปเป็นวัฏจักรหรือวงจรมีภาพที่ 2.1



ภาพที่ 2.1 วงจรประกันคุณภาพภายในสถานศึกษา (Segers and Dochy, 1996)

วงจรประกันคุณภาพภายในสถานศึกษาทั้ง 5 ขั้นตอนได้แก่ ขั้นการเลือกระบบคุณภาพ และกำหนดเป้าหมาย/มาตรฐานการดำเนินงาน/ตัวบ่งชี้ ขั้นการกำหนดยุทธวิธีและการวางแผนการดำเนินงานตามเป้าหมาย ขั้นการดำเนินงานและการติดตามกำกับการดำเนินงานเพื่อตรวจสอบข้อบกพร่องในการดำเนินงาน ขั้นการประเมิน ซึ่งเป็นการรวบรวมข้อมูลเพื่อวัดตัวบ่งชี้การปฏิบัติงาน (performance indicators) และนำมาเทียบกับเกณฑ์ว่าการดำเนินงานได้บรรลุผลตามเกณฑ์ที่กำหนดไว้หรือไม่ อย่างไรก็ตาม ขั้นสุดท้ายคือขั้นการปรับปรุงหรือพัฒนา ซึ่งเป็นการนำผลการประเมินมาใช้เป็นแนวทางพัฒนาการดำเนินงาน/ปรับปรุงการปฏิบัติงานให้ดียิ่งขึ้น จากนั้นนำผลที่ได้จากขั้นการพัฒนามาใช้กำหนดเป้าหมายเพื่อตั้งต้นการดำเนินงานตามวงจรรอบใหม่ต่อไป

การประกันคุณภาพมีภารกิจหลัก 4 ส่วน แต่ละส่วนควรมีหน่วยงานรับผิดชอบที่ชัดเจน (สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา, 2547) ได้แก่ 1) การกำหนดมาตรฐานหรือหลักเกณฑ์ของระบบการประกันคุณภาพเป็นหน้าที่ของส่วนกลาง เช่น กระทรวงศึกษาธิการ 2) การประเมินภายในเป็นหน้าที่ของสถานศึกษาในการพัฒนาระบบการประเมินภายในโดยดึงชุมชนและผู้มีส่วนเกี่ยวข้องเข้ามามีส่วนร่วม 3) การประเมินภายนอกให้

องค์การอิสระในรูปองค์การมหาชนทำหน้าที่ในการประเมินภายนอก และ 4) การนำผลการประเมินไปใช้เพื่อการปรับปรุงเป็นหน้าที่ของฝ่ายบริหารการศึกษา

นอกจากนี้การประกันคุณภาพการศึกษาแย่งแบ่งออกเป็น ระดับการศึกษาขั้นพื้นฐาน ระดับการอาชีวศึกษา และระดับอุดมศึกษา ในที่นี้จะกล่าวถึงเฉพาะการประกันคุณภาพในระดับอุดมศึกษาเท่านั้น

1.1 การประกันคุณภาพระดับอุดมศึกษา

การประกันคุณภาพในระดับอุดมศึกษา เริ่มใช้เมื่อประมาณศตวรรษที่ 19 เมื่อมหาวิทยาลัยเริ่มกำหนดมาตรฐานขั้นต่ำในการรับนักศึกษา และการโอนหน่วยกิตในรายวิชาที่เท่าเทียมกันไปยังมหาวิทยาลัยอื่นๆ (Colbeck et al., 2003 อ้างถึงใน Archwamety, 2006) สหพันธรัฐของอเมริกาให้จัดทำรายชื่อมหาวิทยาลัยที่ได้รับการรับรองย้อนหลังจากปี ค.ศ. 1867 และยังมีรายชื่อของโรงเรียนมัธยมปลายที่ได้รับการอนุมัติจากมหาวิทยาลัยดังกล่าวให้นักเรียนสามารถเข้าเรียนในมหาวิทยาลัยได้ และหลังจากนั้นก็มียุทธศาสตร์หน่วยงานที่ให้การรับรองมหาวิทยาลัยเพิ่มมากขึ้น ตั้งแต่ปลายปี ค.ศ. 1920 ได้แก่ National Commission on Accrediting (NCA) ในปี ค.ศ. 1950 ทำหน้าที่ประสานงานกระบวนการขอการรับรอง ต่อมาเปลี่ยนชื่อเป็น Council on Postsecondary Accreditation (COPA) ในปี ค.ศ. 1975 ซึ่ง COPA ได้ขยายตัวมากขึ้นในปี ค.ศ. 1980 คือ มีตัวแทนอาชีพมากกว่า 70 อาชีพ และสมาคมรับรองในทั่วทั้ง 6 รัฐ (Middle State, New England, North Central, Northwest, Southern, และ Western) นอกจากนี้ในปี ค.ศ. 1993 ได้มีการก่อตั้งสมาคมชื่อ The Association of Specialized and Professional Accreditors (ASPA) ขึ้นแทนหลังจากที่ COPA เลิกดำเนินการ (Fagan and Wells, 2000 อ้างถึงใน Archwamety, 2006) ต่อมาในปี ค.ศ. 1996 ได้ก่อตั้ง The Council for Higher Education Accreditation (CHEA) (Colbeck et al., 2003 อ้างถึงใน Archwamety, 2006) เป็นหน่วยงานที่เกิดจากการรวบรวมตัวกันของมหาวิทยาลัยเอง เพื่อให้การรับรองเรื่องคุณภาพของมหาวิทยาลัย ทำหน้าที่ให้การรับรองในหลายระดับตั้งแต่ระดับสถาบันหรือมหาวิทยาลัย ระดับวิทยาลัย/คณะ ระดับสาขาวิชา เป็นต้น และประเทศอังกฤษได้แสดงความโปร่งใสของมหาวิทยาลัย จึงตั้ง Academic Audit Unit ทำหน้าที่ตรวจสอบกลไกประกันคุณภาพกันภายในในกลุ่มมหาวิทยาลัย นอกจากนี้รัฐบาลยังก่อตั้งหน่วยงาน Higher Education Funding Council เพื่อทำการประเมินคุณภาพ และพิจารณาจัดสรรงบประมาณ สำหรับในประเทศออสเตรเลียและนิวซีแลนด์มีความคล้ายคลึงกัน คือ มีการจัดตั้งหน่วยงานอิสระจากที่ประชุมอธิการบดีของทั้ง 2 ประเทศ คือ Australian Vice-Chancellors Committee (AVCC) กับ New Zealand Vice-Chancellors Committee (NZVCC) ซึ่งเพื่อแยกทำงานประกันคุณภาพจากรัฐ

เพื่อให้เกิดความเป็นอิสระในการดำเนินงานของสถาบันอุดมศึกษา จะเห็นได้ว่าสถาบันอุดมศึกษาจะพยายามรักษาความเป็นอิสระของตนเองไว้ โดยคิดระบบกลไกในการประกันคุณภาพเป็นของตนเอง (อมรวิชัย นาคทรพ, 2541)

สำหรับในประเทศไทยเริ่มจากในปี พ.ศ. 2540 ภายหลังจากมีการประกาศใช้ รัฐธรรมนูญแห่งราชอาณาจักรไทยฉบับใหม่ และได้ประกาศใช้พระราชบัญญัติการศึกษาแห่งชาติ ในปี พ.ศ. 2542 แก้ไขเพิ่มเติมฉบับที่ 2 (พ.ศ. 2545) ได้กำหนดจุดมุ่งหมายและหลักการของการจัดการศึกษาที่มุ่งเน้นคุณภาพและมาตรฐาน โดยกำหนดรายละเอียดในหมวด 6 มาตรฐานและการประกันคุณภาพการศึกษา ประกอบด้วย “ระบบการประกันคุณภาพภายใน” และ “ระบบการประกันคุณภาพภายนอก” เพื่อเป็นกลไกในการรักษาคุณภาพและมาตรฐานของสถาบันอุดมศึกษา

1.2 แนวทางการประกันคุณภาพภายนอก

การประกันคุณภาพภายนอกมีรูปแบบที่หลากหลาย ทั้งที่เป็นหน่วยงานของรัฐ รับผิดชอบ เช่น กรณีของประเทศอังกฤษ หรือนิวซีแลนด์ และหน่วยงานอิสระ ของสหรัฐอเมริกา และหน่วยงานที่ก่อตั้งโดยความร่วมมือของกลุ่มมหาวิทยาลัยเอง หน่วยงานเหล่านี้มีบทบาทและลักษณะการทำงานที่คล้ายคลึงกัน 4 ประการ ดังนี้ 1) เป็นผู้ประสานงาน หรือ วางกรอบการทำงานร่วมกัน มีกลไกกรรมการอำนวยการและกรรมการที่ปรึกษา กรณีหน่วยงานกลางโดยเฉพาะหน่วยงานของรัฐมีแนวโน้มทำหน้าที่กำกับดูแลคุณภาพการศึกษาของสถาบันอุดมศึกษาอย่างใกล้ชิด เพื่อใช้ข้อมูลจากการประกันคุณภาพมาใช้ในการวางแผนและจัดสรรงบประมาณให้สถาบันอุดมศึกษา 2) การให้การสนับสนุนเชิงนโยบายและการเป็นผู้นำทางวิชาการ ซึ่งให้การสนับสนุนทั้งการเงินและการเผยแพร่ข่าวสารเพื่อให้สถาบันอุดมศึกษาสามารถก้าวไปในทิศทางที่ถูกต้อง 3) การทำหน้าที่เฝ้าระวังตรวจสอบ ในกรณีที่เกิดความขัดแย้งต่างๆ ได้แก่ ขัดแย้งกับการเมืองที่เข้ามาตรวจสอบกำกับดำเนินงานของมหาวิทยาลัยทำให้ขาดความเป็นอิสระ 4) การเป็นศูนย์รวบรวมข่าวสาร เกี่ยวกับการดำเนินงานประกันคุณภาพของหน่วยงานต่างๆ เพื่อเป็นแหล่งรวบรวมความรู้และประสบการณ์ร่วมกัน เช่น good practice ในด้านต่างๆ เพื่อเป็นประโยชน์กับสถาบันอื่นๆอีกด้วย

การประเมินคุณภาพภายนอกในประเทศไทย มีสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา มีฐานะเป็นองค์กรมหาชนทำหน้าที่ พัฒนาเกณฑ์ วิธีการประเมินคุณภาพภายนอก และทำการประเมินผลการจัดการศึกษา เพื่อให้มีการตรวจสอบคุณภาพของสถานศึกษา โดยคำนึงถึงจุดมุ่งหมายและหลักการ และแนวการจัดการศึกษาในแต่ละระดับ ตามที่กำหนดไว้ในพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 มาตราที่ 49 และให้มีการประเมินคุณภาพภายนอกของสถานศึกษาทุกแห่งอย่างน้อยหนึ่งครั้งในทุก 5 ปี นับตั้งแต่การ

ประเมินครั้งสุดท้าย และเสนอผลการประเมินต่อหน่วยงานที่เกี่ยวข้องและสาธารณชน ในกรณีที่ผลการประเมินภายนอกของสถานศึกษาใด ไม่ได้ตามมาตรฐานที่กำหนด ให้จัดทำข้อเสนอแนะการปรับปรุง แก้ไขต่อหน่วยงานต้นสังกัด เพื่อให้สถานศึกษาปรับปรุงแก้ไขภายในระยะเวลาที่กำหนด มาตราที่ 51 ดังนั้น การประเมินคุณภาพภายนอกระดับอุดมศึกษา หมายถึง การประเมินคุณภาพการจัดการศึกษา การติดตาม การตรวจสอบคุณภาพและมาตรฐานการศึกษาของสถานศึกษาระดับอุดมศึกษาทุกแห่ง ซึ่งดำเนินการโดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (สมศ.) หรือ ผู้ประเมินภายนอกที่ได้รับการรับรองจาก สมศ. มีจุดมุ่งหมายเพื่อพัฒนาคุณภาพและมาตรฐานการศึกษาของสถาบันการศึกษาระดับอุดมศึกษาให้มีคุณภาพดียิ่งขึ้น (สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา, 2550ก)

การประกันคุณภาพภายนอก เป็นการติดตาม ตรวจสอบ และประเมินคุณภาพและมาตรฐานการศึกษาของสถานศึกษาซึ่งรับผิดชอบดำเนินการโดยหน่วยงานอิสระที่เรียกว่า “สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน)” หรือมีชื่อย่อว่า “สมศ.” สืบเนื่องมาจากพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545 ได้กำหนดให้สถานศึกษาทุกแห่งต้องได้รับการประเมินคุณภาพภายนอกอย่างน้อย 1 ครั้งในรอบ 5 ปี นับตั้งแต่การประเมินครั้งสุดท้าย และเสนอผลการประเมินต่อหน่วยงานที่เกี่ยวข้องและสาธารณชนรับทราบ ซึ่ง สมศ. ได้ดำเนินการประเมินคุณภาพภายนอกรอบแรก (พ.ศ.2544-2548) เสร็จสิ้นไปแล้ว ปัจจุบันอยู่ระหว่างประเมินคุณภาพภายนอกรอบที่สอง (พ.ศ.2549-2553) การประเมินคุณภาพภายนอกรอบที่สองมีความแตกต่างที่สำคัญจากการประเมินรอบแรกสามประการ คือ ประการแรก เป็นการประเมินทั้งในระดับสถาบันและกลุ่มสาขาวิชา ประการที่สอง เป็นการประเมินตามจุดเน้น ของกลุ่มสถาบันว่าสถาบันจะเลือกกำหนดจุดเน้นด้านใด ได้แก่ เน้นการผลิตบัณฑิตและวิจัย เน้นการผลิตบัณฑิตและพัฒนาสังคม เน้นการผลิตบัณฑิตและพัฒนาศิลปะและวัฒนธรรมและเน้นเฉพาะการผลิตบัณฑิต ซึ่งมีตัวบ่งชี้เฉพาะจำนวนหนึ่งที่สถาบันจะต้องถูกประเมินตามจุดเน้น นอกเหนือจากตัวบ่งชี้ร่วม ประการที่สาม การประเมินภายนอกรอบที่สองเป็นการประเมินเพื่อรับรองมาตรฐานคุณภาพทั้งระดับสถาบันและ กลุ่มสาขาวิชา นั้นหมายถึงทุกตัวบ่งชี้จำเป็นต้องมีเกณฑ์กำกับเพื่อตัดสินคุณภาพว่าผ่านการรับรองตามมาตรฐานคุณภาพหรือผ่านการรับรองแบบมีเงื่อนไขหรือไม่ผ่านการรับรอง (สำนักงานคณะกรรมการการอุดมศึกษา, 2550)

หลักในการประเมินคุณภาพ (สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา, 2550) มีการหลักการสำคัญในการประเมิน 5 ประการ ดังต่อไปนี้ 1) เป็นการประเมินเพื่อมุ่งให้มีการพัฒนาคุณภาพการศึกษา ไม่ได้มุ่งเน้นเรื่องการตัดสิน การจับผิด หรือการให้คะแนนให้โทษ

2) ยึดหลักความเที่ยงตรง เป็นธรรม โปร่งใส มีหลักฐานข้อมูลตามสภาพความเป็นจริง (evidence-based) และมีความรับผิดชอบต่อที่ตรวจสอบได้ (accountability) 3) มุ่งเน้นในเรื่องการส่งเสริมและประสานงานในลักษณะกัลยาณมิตรมากกว่าการกำกับควบคุม 4) ส่งเสริมการมีส่วนร่วมในการประเมินคุณภาพและการพัฒนาการจัดการศึกษาจากทุกฝ่ายที่เกี่ยวข้อง 5) มุ่งสร้างความสมดุลระหว่างเสรีภาพทางการศึกษากับจุดมุ่งหมายและหลักการศึกษารัฐธรรมนูญที่กำหนดไว้ในพระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2542 โดยให้ออกภาพเชิงนโยบาย แต่ยังคงมีความหลากหลายในทางปฏิบัติที่สถาบันสามารถกำหนดเป้าหมายเฉพาะและพัฒนาคุณภาพการศึกษาให้เต็มตามศักยภาพของสถาบันและผู้เรียน

2. ความสัมพันธ์ระหว่างมาตรฐานการศึกษาของชาติ กับมาตรฐานการประเมินคุณภาพการศึกษา

การประกันคุณภาพการศึกษา ในระดับอุดมศึกษาของประเทศไทย ประกอบด้วย การประเมินคุณภาพการศึกษา ในระดับอุดมศึกษา ใน 3 ส่วน ดังต่อไปนี้ 1) การประเมินคุณภาพภายใน ตามองค์ประกอบคุณภาพ ตัวบ่งชี้ และเกณฑ์การประเมิน 9 องค์ประกอบ ของสำนักงานคณะกรรมการการอุดมศึกษา 2) การประเมินคุณภาพภายนอก ตามมาตรฐานการประเมินของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา 7 มาตรฐาน และ 3) การประเมินโดยหน่วยงานต้นสังกัด กรณีมหาวิทยาลัยของรัฐ ตามมาตรฐานการประเมินผลการปฏิบัติราชการตามคำรับรองการปฏิบัติราชการของสถาบันอุดมศึกษา 4 มิติ โดยการจัดการศึกษาของไทย ยึดกรอบมาตรฐานการศึกษาของชาติเป็นแนวทางในการจัดการศึกษา และการประเมินคุณภาพการศึกษากว่าคือ มาตรฐานการศึกษาของชาติ มีอุดมการณ์สำคัญของจัดการศึกษา คือ การจัดให้มีการศึกษาตลอดชีวิตและการสร้างสังคมไทยให้เป็นสังคมแห่งการเรียนรู้ โดยกำหนดมาตรฐานไว้ 3 มาตรฐาน ดังต่อไปนี้

มาตรฐานที่ 1 คุณลักษณะของคนไทยที่พึงประสงค์ ทั้งในฐานะพลเมืองและพลโลก

เป้าหมายของการจัดการศึกษาอยู่ที่การพัฒนาคนไทยทุกคนให้เป็น “คนเก่ง คนดี และมีความสุข” โดยมีการพัฒนาที่เหมาะสมกับช่วงวัย พัฒนาคนตามธรรมชาติและเต็มตามศักยภาพ ตรงตามความต้องการ ทั้งในด้านสุขภาพร่างกายและจิตใจ สติปัญญา ความรู้และทักษะคุณธรรมและจิตสำนึกที่พึงประสงค์ และอยู่ในสังคมได้อย่างปกติสุข ประกอบด้วยตัวบ่งชี้หลัก 5 ตัวบ่งชี้

มาตรฐานที่ 2 แนวการจัดการศึกษา

การจัดกระบวนการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญ ผู้เรียนเห็นแบบอย่างที่ดี ได้ฝึกการคิดได้เรียนรู้ จากประสบการณ์ตรงที่หลากหลายตรงตามความต้องการ และมีความสุขในการเรียนรู้ ครู ศึกษานิเทศก์ผู้เรียนเป็นรายบุคคล เตรียมการสอนและใช้สื่อที่ผสมผสานความรู้สากลกับ

ภูมิปัญญาไทย จัดบรรยากาศเอื้อต่อการเรียนรู้จัดหาและพัฒนาแหล่งการเรียนรู้ที่หลากหลาย และพัฒนาความคิดของผู้เรียนอย่างเป็นระบบและสร้างสรรค์

ความสำเร็จของการจัดกระบวนการเรียนรู้ที่เน้นผู้เรียนเป็นสำคัญขึ้นอยู่กับ 1) ปัจจัยด้านบุคคล ได้แก่ ผู้เรียน ครู คณาจารย์ ผู้บริหาร ผู้ปกครอง และสมาชิกชุมชน และ 2) ปัจจัยด้านการบริหาร ได้แก่ หลักการบริหารจัดการและหลักธรรมาภิบาล ประกอบด้วยตัวบ่งชี้หลัก 3 ตัวบ่งชี้

มาตรฐานที่ 3 แนวการสร้างสังคมแห่งการเรียนรู้/สังคมแห่งความรู้

การเรียนรู้ ความรู้ นวัตกรรม สื่อ และเทคโนโลยี เป็นปัจจัยสำคัญของการพัฒนาสู่สังคมแห่งความรู้ การส่งเสริมและสร้างกลไกเพื่อให้คนไทยทุกคนมีโอกาสและทางเลือกที่จะเข้าถึงปัจจัยและการเรียนรู้อย่างต่อเนื่องตลอดชีวิตด้วยรูปแบบและวิธีการที่หลากหลาย โดยการได้รับความร่วมมือจากทุกภาคส่วนของสังคม จะนำมาซึ่งการพัฒนาคุณภาพประสิทธิภาพ และขีดความสามารถของคนไทยในการพัฒนาประเทศ รวมทั้งการเพิ่มศักยภาพการแข่งขันของประเทศ ประกอบด้วยตัวบ่งชี้หลัก 3 ตัวบ่งชี้

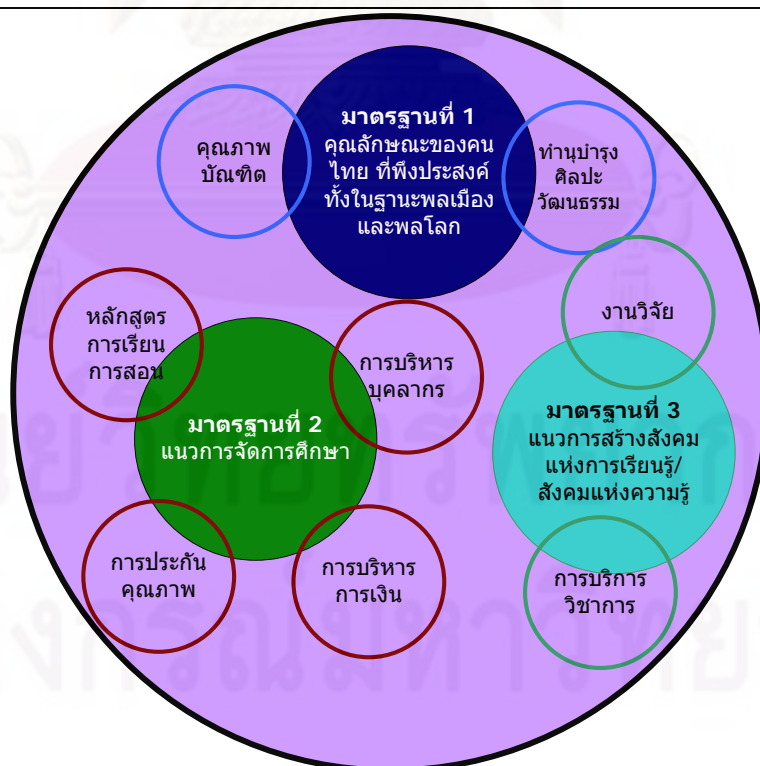
ดังนั้น เมื่อพิจารณามาตรฐานการประเมินคุณภาพการศึกษาของไทย ที่สอดคล้องกับมาตรฐานการศึกษาของชาติ พบว่า มาตรฐานที่ 1 คุณลักษณะของคนไทย ที่พึงประสงค์ ทั้งในฐานะพลเมืองและพลโลก เกี่ยวข้องกับมาตรฐานการประเมินคุณภาพการศึกษา ในด้านคุณภาพบัณฑิตหรือกิจกรรมพัฒนานิสิต และการทำนุบำรุงศิลปวัฒนธรรม และมาตรฐานที่ 2 แนวทางการศึกษา สอดคล้องในด้านหลักสูตรการเรียนการสอน การบริหารจัดการบุคลากร การเงินงบประมาณ และการประกันคุณภาพ ส่วนมาตรฐานที่ 3 แนวการสร้างสังคมแห่งการเรียนรู้/สังคมแห่งความรู้ สอดคล้องกับมาตรฐานการประเมินคุณภาพ เกี่ยวกับการวิจัย และการบริการวิชาการ ดังแสดงในตารางที่ 2.1 และภาพที่ 2.2

ตารางที่ 2.1 ความสัมพันธ์ระหว่างมาตรฐานการศึกษาชาติ กับการประเมินคุณภาพการศึกษา

มาตรฐานการศึกษาของชาติ	สกอ. (2550)	สมศ. (2550)	กพร. (2551)
มาตรฐานที่ 1 คุณลักษณะของคนไทย ที่พึงประสงค์ ทั้งในฐานะพลเมืองและพลโลก	องค์ประกอบที่ 3, 6 คือ กิจกรรมการพัฒนา นิสิตนักศึกษา และการทำนุบำรุง ศิลปะวัฒนธรรม	มาตรฐานที่ 1,4 คือ คุณภาพบัณฑิต และการทำนุบำรุง ศิลปะและวัฒนธรรม	มิติที่ 1 (ตัวชี้วัดที่ 4.1, 4.4) คือ ระดับ ความสำเร็จของการ บรรลุมาตรฐาน สมศ. ได้แก่ ด้านบัณฑิตและ ด้านการทำนุบำรุง ศิลปะและวัฒนธรรม

ตารางที่ 2.1 (ต่อ)

มาตรฐานการศึกษาของชาติ	สกอ. (2550)	สมศ. (2550)	กพร. (2551)
มาตรฐานที่ 2 แนวการจัดการศึกษา	องค์ประกอบที่ 1, 2,7,8,9 คือ ปรัชญา ปณิธาน วัตถุประสงค์ และแผนการดำเนินการ , การเรียนการสอน, การบริหารและการจัดการ, การเงินและงบประมาณ , และระบบและกลไก การประกันคุณภาพ	มาตรฐานที่ 5,6,7 คือ การพัฒนาสถาบัน และบุคลากร, หลักสูตร และการเรียนการสอน, ระบบการประกัน คุณภาพ	มิติที่1,2,3,4 คือ ด้าน ประสิทธิภาพ, คุณภาพ, ประสิทธิภาพของการ ปฏิบัติราชการ, และ การพัฒนาสถาบัน
มาตรฐานที่ 3 แนวการสร้างสังคมแห่งการเรียนรู้/สังคมแห่งความรู้	องค์ประกอบที่ 4, 5 คือ การวิจัย, การบริการ วิชาการแก่สังคม	มาตรฐานที่ 2, 3 คือ งานวิจัยและงานสร้าง สรรค์ และการบริการ วิชาการ	มิติที่ 1 (ตัวชี้วัดที่ 4.2, 4.3) คือ ระดับ ความสำเร็จของการ บรรลุมาตรฐาน สมศ. ได้แก่ การวิจัยและงาน สร้างสรรค์ และการ บริการวิชาการ



ภาพที่ 2.2 ความสัมพันธ์ของมาตรฐานการศึกษาของชาติกับการประเมินคุณภาพการศึกษา

การประกันคุณภาพเป็นเครื่องมือสำคัญที่ช่วยให้ผู้บริหารทราบว่าการดำเนินงานของสถานศึกษามีความโปร่งใส และตรวจสอบได้ รู้จุดแข็ง จุดที่ควรพัฒนาและสามารถกำหนดแนวทางการพัฒนาสถานศึกษาได้อย่างมีประสิทธิภาพ การประเมินคุณภาพภายนอกมีการดำเนินการโดยสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) มีจุดมุ่งหมายเพื่อพัฒนาคุณภาพและมาตรฐานการศึกษาของสถาบันการศึกษาให้มีคุณภาพดียิ่งขึ้น สำหรับองค์ประกอบการประเมินคุณภาพภายในของสำนักงานคณะกรรมการการอุดมศึกษา 9 องค์ประกอบ 44 ตัวบ่งชี้ และตัวบ่งชี้การประเมินภายนอกของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา(องค์การมหาชน) ประกอบด้วย 7 มาตรฐาน 48 ตัวบ่งชี้ นอกจากนี้มีมาตรฐานการประเมินผลการปฏิบัติราชการตามคำรับรองการปฏิบัติราชการของสถาบันอุดมศึกษา สำหรับสถาบันอุดมศึกษาของรัฐ ประกอบด้วย 4 มิติ คือ มิติด้านประสิทธิผล มิติด้านคุณภาพ มิติด้านประสิทธิภาพของการปฏิบัติราชการ และมิติด้านการพัฒนาสถาบัน 23 ตัวบ่งชี้ และมีตัวบ่งชี้ตัวเดียวกันรวม 12 ตัวบ่งชี้ ส่วนตัวบ่งชี้อื่นๆ มีความสอดคล้องหรือใช้ข้อมูลร่วมกันอยู่ด้วย

3. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการประกันคุณภาพและการประเมินคุณภาพ จำแนกได้ 3 กลุ่ม คือ 3.1) กลุ่มงานวิจัยที่เกี่ยวข้องกับการประกันคุณภาพ พัฒนาโมเดลการประกันคุณภาพ ได้แก่ งานวิจัยของ รุ่งอรุณ สุเวชวัฒนกุล (2543) 3.2) งานวิจัยที่เกี่ยวกับวิธีการประเมินและกระบวนการประเมิน ได้แก่ งานวิจัยของ Russon และ Koehly (1995), นางลักษณีย์ วิรัชชัย และ วรณี เจตจำนงนุช (2548), และ Bornmann, Mittag และ Daniel (2006) 3.3) การใช้ผลการประเมิน ได้แก่ วราภรณ์ บุญเยี่ยม (2546) และ รติกร ชันติยานันท์ (2548) มีรายละเอียดดังนี้

3.1 งานวิจัยที่เกี่ยวข้องกับการพัฒนาโมเดลการประกันคุณภาพการอุดมศึกษา คณะครุศาสตร์ ศึกษาศาสตร์ สังกัดทบวงมหาวิทยาลัย (รุ่งอรุณ สุเวชวัฒนกุล, 2543) ทำการสร้างและตรวจสอบโมเดลผลการประกันคุณภาพการอุดมศึกษา คณะครุศาสตร์ ศึกษาศาสตร์ สังกัดทบวงมหาวิทยาลัย เครื่องมือที่ใช้เป็นแบบสอบถาม โดยสอบถามอาจารย์ และตรวจสอบโมเดลการวิจัยกับข้อมูลเชิงประจักษ์ ด้วยโปรแกรมลิสเรล พบว่า โมเดลสามารถอธิบายความแปรปรวนของผลการประกันคุณภาพการอุดมศึกษาได้ร้อยละ 95 สอดคล้องกับข้อมูลเชิงประจักษ์ ตัวแปรที่ค่าอิทธิพลรวมส่งผลต่อผลการประกันคุณภาพการอุดมศึกษาอย่างมีนัยสำคัญ เรียงจากน้อยไปหามาก คือ ระบบและกลไกการประกันคุณภาพ การบริหารจัดการ การเรียนการสอน การวิจัย การเงินและงบประมาณ การพัฒนานิสิต การบริการวิชาการแก่สังคม ปรัชญา ปณิธาน วัตถุประสงค์และแผนงาน และการทำงานบูรณาการศิลปวัฒนธรรม ตามลำดับ

3.2 งานวิจัยที่ใช้วิธีการประเมินในเชิงคุณภาพและเชิงปริมาณ โดยการสร้างมาตรวัดผลกระทบที่น่าเชื่อถือของรายงานการประเมิน จากงานวิจัยของ Russon และ Koehly (1995) พัฒนามาตรวัดโดยนักจิตวิทยาใช้มาตรวัดแบบ Likert 28 ข้อ เพื่อวัดมุมมอง การเชื่อมโยง ความเข้าใจ การอ้างอิงเกี่ยวกับความรู้ การเรียนรู้ ทักษะคิด และการตัดสินใจ กลุ่มตัวอย่างเชิงสำรวจ 2 กลุ่ม และเชิงยืนยัน 2 กลุ่ม ซึ่งมีค่าความเที่ยงมากกว่าหรือเท่ากับ .80 และ confirmatory factor analysis มีค่า coefficient ระหว่าง .70-.80 มีความเที่ยงในระดับดีมาก เครื่องมือนี้สามารถช่วยให้ผู้วิจัยและผู้ทำการประมาณค่าผลกระทบของรายงาน ภายใต้การตัดสินใจตามวิธีวิทยาการประเมินที่แตกต่างกัน หากเป็นการประเมินเชิงปริมาณ พิจารณาถึงตารางความถี่ กราฟ และตารางข้อมูลสถิติ ส่วนการประเมินเชิงคุณภาพ พิจารณาถึงรายละเอียดการบรรยาย ข้อมูลจากการสัมภาษณ์ และผลจากการวิเคราะห์เอกสาร และนอกจากนี้งานวิจัยที่เกี่ยวกับวิธีการประเมินและกระบวนการประเมิน ได้แก่ งานวิจัยของนางลักษณะ วิรัชชัย และ วรณี เจตจำนงนุช (2548) ทำการสังเคราะห์งานประเมินจากรายงานผลการประเมินคุณภาพภายนอก ระดับอุดมศึกษา ตามหลักการประเมินอภิมาน (meta-evaluation) พบว่า ลักษณะของรายงานผลการประเมินคุณภาพภายนอกสถาบันอุดมศึกษามีลักษณะความแตกต่างระหว่างสถาบันค่อนข้างมาก รายงานส่วนใหญ่มีคุณภาพตามมาตรฐานการประเมินอภิมาน ลักษณะและวิธีการประเมินคุณภาพภายนอกส่วนใหญ่ตรงตามข้อกำหนดการประเมินคุณภาพภายนอกของสมศ. ผลการประเมินคุณภาพสถาบันอุดมศึกษา ตามมาตรฐานและตัวบ่งชี้การประเมินคุณภาพภายนอก ที่ สมศ.กำหนด พบว่า คุณภาพสถาบันอุดมศึกษามีความแตกต่างกัน ตามรูปแบบการประเมินคุณภาพภายนอกระดับอุดมศึกษาของ สมศ. นักประเมินภายนอกมีคุณสมบัติตามข้อกำหนดของ สมศ. เมื่อพิจารณาแยกตามสังกัดของสถาบันอุดมศึกษา พบว่าคณะนักประเมินภายนอกของมหาวิทยาลัยภาครัฐ มีวุฒิปริญญาเอก และมีตำแหน่งวิชาการระดับศาสตราจารย์ มากกว่าจำนวนนักประเมินภายนอกของมหาวิทยาลัยเอกชนและมหาวิทยาลัยราชภัฏอยู่เล็กน้อย และงานวิจัยการประเมินอภิมาน โดย Bormann, Mittag และ Daniel (2006) ใช้การประเมินอภิมานกระบวนการหลายขั้นตอนใช้ในการประเมินการเรียนการสอนที่ได้รับการยอมรับและน่าเชื่อถือในประเทศเยอรมัน ซึ่งผลการวิจัย พบว่า การศึกษาถึงความเข้าใจและเป็นตัวแทนของกระบวนการสำหรับการประเมินการสอนและการเรียนในเยอรมัน ซึ่งผู้เข้าร่วมในการประเมินต่างพอใจในภาพรวมของกระบวนการแบบหลายขั้นตอน และเชื่อในจุดมุ่งหมายของการประกันคุณภาพและการปรับปรุงเพื่อให้บรรลุผลสำเร็จ ข้อเสนอแนะสำหรับการปรับปรุงเป้าหมายกระบวนการรายบุคคล ตัวอย่างเช่น องค์ประกอบของคณะกรรมการประเมินมีพื้นฐานที่เข้าใจถึงการปฏิบัติในการประเมินคุณภาพของกระบวนการอย่างถูกต้องสำหรับประกันคุณภาพและการปรับปรุง

3.3 การใช้ผลการประเมิน ได้แก่ การใช้ผลการประเมินในเชิงแนวคิด เชิงตรวจสอบยืนยัน เชิงสัญลักษณ์ และเชิงปฏิบัติ ในงานวิจัยของ วราภรณ์ บุญเยี่ยม (2546) และรติกร ชันติยานันท์ (2548) ซึ่งได้ศึกษาเพิ่มเติมถึงปัจจัยที่ส่งผลกระทบต่อการใช้ผลการประเมิน โดยผลการวิจัยของวราภรณ์ บุญเยี่ยม (2546) พบว่า สถานศึกษาส่วนใหญ่ใช้ผลการประเมินภายนอกในเชิงแนวคิด เชิงตรวจสอบยืนยัน และเชิงสัญลักษณ์อย่างครบถ้วน มากกว่า ใช้ผลการประเมินภายนอกในเชิงแนวคิด ตรวจสอบยืนยัน และสัญลักษณ์เพียงบางส่วน และไม่ใช้ผลการประเมินภายนอกในเชิงแนวคิด ตรวจสอบยืนยัน และสัญลักษณ์ และสถานศึกษาใช้ผลการประเมินภายนอกเชิงปฏิบัติเพียงบางส่วน มากกว่า ใช้ผลการประเมินภายนอกในเชิงปฏิบัติอย่างครบถ้วน และไม่ใช้ผลการประเมินภายนอกในเชิงปฏิบัติ และยังพบว่า ปัจจัยที่ส่งผลกระทบต่อการใช้ผลการประเมินภายนอกเพื่อพัฒนาคุณภาพของสถานศึกษาขั้นพื้นฐานทั้ง 4 รูปแบบอย่างมีนัยสำคัญทางสถิติ คือ สถานศึกษาที่มีผลการประเมินภายนอกในระดับควรปรับปรุง และทัศนคติต่อการประเมินภายนอกของผู้บริหารสถานศึกษาตามลำดับ นอกจากนี้สถานศึกษาที่มีผลการประเมินภายนอกในระดับพอใช้ มีอิทธิพลต่อการใช้ผลการประเมินภายนอกอย่างครบถ้วนในเชิงแนวคิด ในเชิงตรวจสอบยืนยัน และในเชิงสัญลักษณ์ เปลี่ยนแปลงเพิ่มขึ้น และทักษะการสื่อสารของผู้ประเมินภายนอกมีอิทธิพลต่อการใช้ผลการประเมินภายนอกในเชิงปฏิบัติอย่างครบถ้วนเปลี่ยนแปลงเพิ่มขึ้น แต่งานวิจัยของรติกร ชันติยานันท์ (2548) พบว่า โรงเรียนที่จัดการศึกษาขั้นพื้นฐานที่ผ่านการประเมินคุณภาพภายนอกในรอบแรกจากสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษานั้น โรงเรียนส่วนใหญ่มีการใช้ผลการประเมินอยู่ในระดับดีทั้ง 4 รูปแบบ คือ เชิงแนวคิด เชิงตรวจสอบยืนยัน เชิงสัญลักษณ์ และเชิงปฏิบัติ ยกเว้นโรงเรียนขนาดเล็กที่มีผลการประเมินภายนอกควรปรับปรุงการใช้ผลการประเมินภายในเชิงตรวจสอบยืนยันในระดับพอใช้ และคุณภาพการประเมินภายในของโรงเรียนส่วนใหญ่อยู่ในระดับดี ทั้ง 4 ด้าน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสม และด้านความถูกต้อง ยกเว้นโรงเรียนขนาดเล็กที่มีผลการประเมินภายนอกในระดับควรปรับปรุงมีคุณภาพการประเมินภายในด้านอรรถประโยชน์และความถูกต้องระดับพอใช้

ตอนที่ 2 มาตรฐานการประเมิน เกณฑ์การประเมิน

แนวคิดทฤษฎีเกี่ยวกับมาตรฐานการประเมิน เกณฑ์การประเมินในส่วนนี้แบ่งออกเป็น 3 หัวข้อ ได้แก่ 1) ความหมายของมาตรฐาน และเกณฑ์การประเมิน 2) มาตรฐานการประเมินทางการศึกษา และ 3) งานวิจัยที่เกี่ยวข้อง มีรายละเอียดดังต่อไปนี้

1. ความหมายของมาตรฐาน และเกณฑ์การประเมิน

มาตรฐานและเกณฑ์การประเมิน เริ่มมาจากในระยะแรกเป็นวัดผลในโรงเรียน เนื่องจากการประเมินแฝงอยู่ในศาสตร์ทางการวัด จึงมักเรียกว่า การวัดและการประเมิน ซึ่ง Thomdike และ Hagen ได้กล่าวว่าการประเมินเป็นกระบวนการที่มีความสัมพันธ์อย่างใกล้ชิดกับการวัดเพียงแต่ว่าการประเมินมีความเฉพาะเจาะจงมากกว่า แต่ต้องอาศัยการวินิจฉัยและการวัดที่ดีมีคุณภาพเป็นที่ยอมรับเป็นพื้นฐาน สำหรับมาตรฐาน หมายถึง คุณลักษณะหรือระดับที่ถือเป็นคุณภาพ ความสำเร็จ หรือความเหมาะสมอันเป็นที่ยอมรับกันทางวิชาชีพ (ศิริชัย กาญจนวาสี, 2547ก) และท.เทียนประทีป (2546) กล่าวถึงมาตรฐานว่า เป็นเอกสารที่จัดทำขึ้นจากการเห็นพ้องต้องกัน และได้รับความเห็นชอบจากองค์กรอันเป็นที่ยอมรับกันทั่วไป เอกสารดังกล่าววางกฎระเบียบแนวทางปฏิบัติหรือลักษณะเฉพาะกิจกรรมนั้นๆ เพื่อให้เป็นหลักเกณฑ์ที่ใช้กันทั่วไปจนเป็นปกติวิสัย โดยมุ่งให้บรรลุถึงความสำเร็จสูงสุดตามข้อกำหนดที่วางไว้ ซึ่งการกำหนดมาตรฐาน (standard setting) ในการทดสอบนั้น เป็นกระบวนการทางจิตมิติ (psychometric) ที่เป็นการกำหนดกฎเกณฑ์ในการตัดสินผลการประเมิน ตามระดับคะแนนผลการทดสอบโดยอาจจำแนกผู้ถูกประเมินเป็นกลุ่มตั้งแต่สองกลุ่มขึ้นไป ทั้งนี้กฎเกณฑ์ในการตัดสินผลการประเมินต้องระบุถึงสมรรถภาพขั้นต่ำสุดที่ยอมรับได้ ในที่นี้ คะแนนจุดตัดที่ใช้ในการตัดสินผลการประเมินเรียกว่า *คะแนนจุดตัด (cut score)* หรือ *เกณฑ์การประเมิน (evaluation criteria)* ซึ่งหมายถึงคุณลักษณะหรือระดับที่ถือว่าเป็นคุณภาพ ความสำเร็จหรือความเหมาะสมของทรัพยากร การดำเนินงานหรือผลการดำเนินงาน นักประเมินกำหนดมาตรฐานการประเมิน ในรูปข้อความแสดงถึงสมรรถภาพของผู้เรียนแต่ละมิติที่ประเมิน และมีเกณฑ์การประเมินหรือคะแนนจุดตัด เป็นกฎสำหรับตัดสินผลการประเมิน โดยสามารถจำแนกเกณฑ์ออกเป็น 2 ประเภท คือ *เกณฑ์สัมพัทธ์ (relative criterion)* และ *เกณฑ์สัมบูรณ์ (absolute criterion)* โดยเกณฑ์สัมพัทธ์เป็นเกณฑ์ที่พัฒนามาจากประสบการณ์ในการเปรียบเทียบผลระหว่างโครงการ หรือเปรียบเทียบกับปกติวิสัย (norm) ของการจัดโครงการโดยทั่วไป ส่วนเกณฑ์สัมบูรณ์เป็นเกณฑ์ที่พัฒนามาจากหลักเหตุผลเกี่ยวกับมาตรฐานของสิ่งนั้น หรือความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับมาตรฐานอันเป็นที่ยอมรับทางวิชาชีพ หรือคุณภาพของสิ่งนั้นอันเป็นที่ยอมรับของผู้เกี่ยวข้องอื่นๆ (ศิริชัย กาญจนวาสี, 2547ก) ด้วยเหตุนี้นักประเมินหลายคน จึงนิยมเรียกทั้งมาตรฐานและเกณฑ์การประเมินว่า **เกณฑ์มาตรฐานการประเมิน (evaluation standard criteria)**

2. มาตรฐานการประเมินทางการศึกษา

มาตรฐานการประเมินทางการศึกษา ประกอบด้วย 3 หัวข้อ ได้แก่ 2.1) มาตรฐานการประเมิน ของ Joint Committee on Standard for Educational Evaluation ,1994 2.2)

มาตรฐานการประเมินที่เกี่ยวข้องกับผู้ประเมิน 2.3 มาตรฐานและตัวบ่งชี้ในการประเมินอภิमान
รายงานการประเมินคุณภาพภายนอกในระดับอุดมศึกษา รายละเอียดดังต่อไปนี้

2.1 มาตรฐานการประเมิน ของ Joint Committee on Standard for Educational Evaluation ,1994 The Program Evaluation Standards ซึ่ง Sander (1994; สมหวัง พิธิยานุวัฒน์, 2549 อ้างถึงใน Sander, 1994) กล่าวว่า มาตรฐานการประเมินเริ่มต้นตั้งแต่นั้นปี ค.ศ. 1974 ริเริ่มโดยสมาคมวิจัยการศึกษาของสหรัฐอเมริกา (The American Educational Research Association) สมาคมจิตวิทยาของสหรัฐอเมริกา (The American Psychological Association) และสภาระดับชาติว่าด้วยการวัดทางการศึกษา (The National Council on Measurement in Education) ได้แต่งตั้งคณะกรรมการเพื่อปรับปรุงมาตรฐานของแบบสอบทางการศึกษาและจิตวิทยา ฉบับปี ค.ศ. 1966 ได้มีการพิจารณาเพิ่มเรื่องการประเมินค่า (evaluation) ในมาตรฐานของแบบสอบ จึงได้แต่งตั้งคณะกรรมการเพิ่มขึ้นอีกชุดหนึ่ง เพื่อกำหนดมาตรฐานของการใช้แบบสอบในการประเมินทางการศึกษา ที่ได้รับความเห็นชอบจากทั้ง 3 สมาคม และผู้แทนจากสมาคมๆละ 3 คน คณะกรรมการเห็นว่าควรมีโครงการจัดทำมาตรฐานการประเมินทางการศึกษาที่ใช้ในการประเมินทั่วไป ไม่ใช่เฉพาะการใช้แบบสอบในการประเมินทางการศึกษาเท่านั้น จึงได้เพิ่มคณะกรรมการจาก 3 สมาคมเป็น 12 สมาคม เพื่อกำหนดมาตรฐานสำหรับการประเมินทางการศึกษาในปี ค.ศ. 1981 โดยมี Stufflebeam เป็นประธาน และได้จัดพิมพ์มาตรฐานสำหรับการประเมินโปรแกรม โครงการและวัสดุอุปกรณ์ ไม่รวมถึงมาตรฐานการประเมินสถาบัน และต่อมามีการกำหนดมาตรฐานการประเมินบุคลากรตีพิมพ์ในปี ค.ศ. 1988 หลังจากทีคณะกรรมการได้ทดลองใช้มาตรฐานการประเมินที่กำหนดขึ้นก่อนนำไปใช้จริง มาตรฐานการประเมินได้รับการรับรองจากสถาบันแห่งชาติว่าด้วยมาตรฐานของประเทศสหรัฐอเมริกา (the American National Standards Institute: ANSI) ในปี ค.ศ. 1989 และมาตรฐานสำหรับการประเมินทางการศึกษาได้รับการรับรองและเผยแพร่ไปทั่วโลก มาตรฐานที่กำหนดขึ้นประกอบด้วย 4 หมวด คือ มาตรฐานด้านอรรถประโยชน์ (utility Standards) ประกอบด้วย 8 มาตรฐานย่อย มาตรฐานด้านความเป็นไปได้ (feasibility standards) ประกอบด้วย 3 มาตรฐานย่อย มาตรฐานด้านความเหมาะสม (propriety standards) ประกอบด้วย 8 มาตรฐานย่อย และมาตรฐานด้านความถูกต้อง (accuracy standards) ประกอบด้วย 11 มาตรฐานย่อย รวมทั้งสิ้น 30 มาตรฐานย่อย

การนำมาตรฐานการประเมินไปประยุกต์ใช้สอดคล้องกับบริบทของประเทศต่างๆ ในหลายประเทศ ได้แก่ มาตรฐานการประเมินในบริบทของประเทศสวิตเซอร์แลนด์ เรียกว่า SEVAL's Evaluation Standards พัฒนาโดยสมาคมการประเมินของประเทศสวิตเซอร์แลนด์ (Swiss Evaluation Society) นำโดย Widmer จากมหาวิทยาลัยซูริค (Zurich University) กำหนดมาตรฐาน

การประเมินหลัก 4 มาตรฐาน และ 28 มาตรฐานย่อย ซึ่งละเว้นมาตรฐานย่อย P1, P8 เพิ่ม U2 Clarifying the Objectives of the Evaluation รวมมาตรฐานย่อย A5 , A6 เข้าด้วยกันเป็น A5 Valid and Reliable Information และรวมมาตรฐานย่อย A8, A9 เป็น A7 Analysis of Qualitative and Quantitative Information (Laubli Loud, 2004) มาตรฐานการประเมินในบริบทของประเทศเยอรมัน The DeGEval's Standards ที่พัฒนาโดยสมาคมการประเมินของประเทศเยอรมัน (German Evaluation Society) นำโดย Beywl (2003) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 25 มาตรฐานย่อย โดยละเว้นมาตรฐานย่อย P1, P7, และ P8 เพิ่มมาตรฐานย่อยเพิ่ม U2 Clarification of the Purposes of the Evaluation รวมมาตรฐานย่อย A5 , A6 เข้าด้วยกันเป็น A5 Valid and Reliable Information และรวมมาตรฐานย่อย A8, A9 เป็น A7 Analysis of Qualitative and Quantitative Information และมาตรฐานการประเมินในบริบทของประเทศแอฟริกา The African Evaluation guidelines 2002 (AEG) ที่พัฒนาขึ้นโดยสมาคมการประเมินของประเทศแอฟริกา (African Evaluation Association) นำโดย Patel (2002) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 30 มาตรฐานย่อย โดยมีการปรับปรุงในมาตรฐานย่อย 14 มาตรฐาน ที่เทียบ AEG ในมาตรฐาน U1, U4, U6, F2, F3, P2, P3, P4, P6, P7, A1, A2, A4, A5 ตามลำดับ จึงนำได้นำมามาตรฐานของแต่ละประเทศ มาเปรียบเทียบกับ The Program Evaluation Standards ดังรายละเอียดดังตารางที่ 2.2

ตารางที่ 2.2 การเปรียบเทียบการประยุกต์ใช้ The Program Evaluation Standards ในประเทศแอฟริกา (Africa) เยอรมัน (German) และสวิสเซอร์แลนด์ (Swiss)

Program Evaluation Standards	AEG (2002) Africa	DeGEval (2003) German	SEVAL (2004) Swiss
ด้านอรรถประโยชน์ (Utility)			
U1 การระบุผู้มีส่วนได้ส่วนเสีย (Stakeholder Identification)	✓ AEG	✓	U1 Identifying Stakeholders
U2 ความน่าเชื่อถือของผู้ประเมิน (Evaluator Credibility)	✓	U3 Evaluator Credibility and Competence	U3 Credibility
U3 ขอบเขตและการเลือกการเก็บรวบรวมข้อมูล (Information Scope and Selection)	✓	✓ U4	U4 Scope and Selection of Information
U4 การระบุคุณค่า (Values Identification)	✓ AEG	U5 Transparency of Values	U5 Transparency of Value Judgments
U5 ความชัดเจนของรายงานการประเมิน (Report Clarity)	✓	U6 Report Comprehensiveness and Clarity	U6 Comprehensiveness and Clarity in Reporting

ตารางที่ 2.2 (ต่อ)

Program Evaluation Standards	AEG (2002)	DeGEval (2003)	SEVAL (2004)
	Africa	German	Swiss
U6 การรายงานภายในเวลาที่กำหนด และการเผยแพร่รายงาน (Report Timeliness and Dissemination)	✓ AEG	U7 Evaluation Timeliness	U7 Timely Reporting
U7 ผลกระทบของงานประเมิน (Evaluation Impact)	✓	U8 Evaluation Utilization and Use	✓ U8
ด้านความเป็นไปได้(Feasibility)			
F1 กระบวนการปฏิบัติ (Practical Procedures)	✓	F1 Appropriate Procedures	✓
F2 การเป็นที่ยอมรับทางการเมือง (Political Viability)	✓ AEG	F2 Diplomatic Conduct	F2 Anticipating Political Viability
F3 ประสิทธิภาพของต้นทุนการประเมิน (Cost Effectiveness)	✓ AEG	F3 Evaluation Efficiency	✓
ด้านความเหมาะสม (Propriety)			
P1 การมุ่งเน้นความสำคัญของการให้บริการ (Service Orientation)	✓	-	-
P2 การทำข้อตกลงอย่างเป็นทางการ (Formal Agreements)	✓ AEG	✓ P1	P1 Formal Written Agreement
P3 การปกป้องสิทธิมนุษยชนของตัวอย่าง (Rights of Human Subjects)	✓ AEG	P2 Protection of Individual Rights	P2 Ensuring Individual Rights and Well-Being
P4 การมีปฏิสัมพันธ์กับผู้ที่เกี่ยวข้อง (Human Interactions)	✓ AEG	-	P3 Respecting Human Dignity
P5 การประเมินที่สมบูรณ์และยุติธรรม (Complete and Fair Assessment)	✓	P3 Complete and Fair Investigation	P4 Complete and Balanced Assessment
P6 การเปิดเผยผลการประเมิน (Disclosure of Findings)	✓ AEG	✓ P5	P5 Making Finding Available
P7 ผลประโยชน์ทับซ้อน (Conflict of Interest)	✓ AEG	-	P6 Declaring Conflicts of Interest
P8 ความรับผิดชอบในเรื่องการเงิน (Fiscal Responsibility)	✓	-	-

ตารางที่ 2.2 (ต่อ)

Program Evaluation Standards	AEG (2002) Africa	DeGEval (2003) German	SEVAL (2004) Swiss
ด้านความถูกต้อง (Accuracy)			
A1การจัดการระบบเอกสารของโครงการ (Program Documentation)	✓ AEG	A1 Description of the Evaluand	A1 Precise Description of the Object of Evaluation
A2 การวิเคราะห์บริบทของโครงการ (Context Analysis)	✓ AEG	✓	A2 Analyzing the Context
A3 การบรรยายวัตถุประสงค์และกระบวนการประเมิน (Described Purposes and Procedures)	✓	✓	A3 precise Description of Goals, Questions, and Procedures
A4 การใช้แหล่งสารสนเทศที่น่าเชื่อถือ (Defensible Information Sources)	✓ AEG	A4 Disclosure of Information Sources	A4 Trustworthy Sources of Information
A5 สารสนเทศที่ได้มีความตรง (Valid Information)	✓ AEG	A5 Valid and Reliable Information	A5 Valid and Reliable Information
A6 สารสนเทศที่ได้มีความเที่ยง (Reliable Information)	✓		
A7 ความเป็นระบบของสารสนเทศ (Systematic Information)	✓	A6 Systematic Data Review	A6 Systematic Checking for Errors
A8 การวิเคราะห์สารสนเทศเชิงปริมาณอย่างมีประสิทธิภาพ (Analysis of Quantitative Information)	✓	A7 Analysis of Qualitative and Quantitative Information	A7 Qualitative and Quantitative Analysis
A9 การวิเคราะห์สารสนเทศเชิงคุณภาพอย่างมีประสิทธิภาพ (Analysis of Qualitative Information)	✓		
A10 การลงข้อสรุปที่มีเหตุผลเหมาะสมถูกต้อง (Justified Conclusions)	✓	✓ A8	A8 Substantiated Conclusions
A11 การรายงานผลอย่างเป็นไม่มีอคติ (Impartial Reporting)	✓	P4 Unbiased Conduct and Reporting	A9 Neutral Reporting
A12 การประเมินอภิมาน (Metaevaluation)	✓	A9 Meta-Evaluation	✓ A10
Total	30	25	27

หมายเหตุ

1. German เพิ่ม U2 Clarification of the Purposes of the Evaluation
2. Swiss เพิ่ม U2 Clarifying the Objectives of the Evaluation

นอกจากนี้ยังมีมาตรฐานการประเมินที่สอดคล้องกับ The Joint Committee on Standard for Educational Evaluation (1994) ที่เน้นความสำคัญของนักประเมิน ในด้านสมรรถนะ ความรู้ความสามารถพื้นฐาน คุณลักษณะของนักประเมิน และลักษณะการจัดการประเมิน อีกด้วย ได้แก่ Guiding Principles for Evaluators ของ American Evaluation Association [AEA] (1995) Essential Skills Series in Evaluation ของ Canadian Evaluation Society (1999) และงานวิจัยของ Ghere & et al. (2006) เพื่อพัฒนา Essential Competencies for Program Evaluators เป็น มาตรฐานที่สอดคล้องกับมาตรฐานย่อย ในมาตรฐานการประเมิน ได้แก่ A7, U2, P3, P4 และ มาตรฐานย่อยอื่นๆที่เกี่ยวข้องกับตัวนักประเมิน ซึ่งสามารถนำมาประยุกต์ใช้ในการประเมิน คุณลักษณะของนักประเมินที่ดี

2.2 มาตรฐานการประเมินที่เกี่ยวข้องกับผู้ประเมิน

มาตรฐานการประเมินที่เกี่ยวข้องกับผู้ประเมิน ประกอบด้วย มาตรฐานของ สมาคมประเมินของสหรัฐอเมริกา Guiding Principles for Evaluators ของ American Evaluation Association [AEA] (2003) ทักษะที่จำเป็นในการประเมินของประเทศแคนาดา Essential Skills Series in Evaluation ของ Canadian Evaluation Society (1999) และสมรรถนะที่จำเป็นสำหรับผู้ประเมิน Essential Competencies for Program Evaluators จากงานวิจัยของ Ghere และคณะ (2006) ซึ่งสามารถสรุปประเด็นที่เกี่ยวข้องกับมาตรฐานการประเมินได้ 5 ประเด็นดังต่อไปนี้ 1) การ ดำเนินการสืบค้นอย่างเป็นระบบ (systematic inquiry) เป็นการประเมินโดยเน้นเทคนิคการประเมิน ตั้งแต่การออกแบบ การวัด การวิเคราะห์ข้อมูล การแปลผล โดยใช้วิธีการเชิงปริมาณและเชิงคุณภาพ ซึ่งระบุแหล่งข้อมูล การเก็บรวบรวมข้อมูล ประเมินความเที่ยง ความตรงของข้อมูล มีเหตุผลเพียงพอ ในการตัดสินใจผลการประเมิน บันทึกจุดแข็ง และข้อจำกัดของการประเมิน 2) สมรรถนะในการ ปฏิบัติงานของนักประเมิน (competence) หมายถึง ผู้ประเมินต้องมีสมรรถนะในการปฏิบัติงาน ได้แก่ ความรู้ ความสามารถ ทักษะการสื่อสารด้วยวาจา การฟัง การเขียน และประสบการณ์ที่เหมาะสม เข้าใจวิธีการประเมิน รูปแบบคำถามและการออกแบบการประเมิน พัฒนาความเชี่ยวชาญในการประเมิน เพื่อสามารถดำเนินการประเมินได้บรรลุเป้าหมาย 3) ความซื่อสัตย์ของผู้ ประเมิน (integrity/honesty) หมายถึง ผู้ประเมินมีความซื่อสัตย์ต่อกระบวนการประเมิน มีการระบุ กลุ่มลูกค้า ออกแบบการประเมินตั้งแต่เริ่มต้น จนเสร็จสิ้น สื่อสารกับผู้ถูกประเมินตลอดขั้นตอน การประเมิน แสดงเหตุผลเกี่ยวกับต้นทุน ใช้เทคโนโลยีให้เหมาะสม จัดการประเมินโดยวิธีที่รัดกุม ในช่วงเวลาที่กำหนด 4) การเคารพต่อสิทธิ ศักดิ์ศรีส่วนบุคคล (respect for people) หมายถึง การ คำนึงถึงความปลอดภัย เกียรติยศ และคุณค่าในตนเองของผู้ให้ข้อมูล ผู้มีส่วนร่วม และลูกค้า

โดยต้องยึดมาตรฐานและคุณธรรมของอาชีพ ความเสี่ยงภัย อันตรายและความทุกข์ใจที่จะเกิดกับผู้มีส่วนร่วมในการประเมิน กำหนดวิธีการประเมินตามความต้องการของผู้รับผลการประเมิน โดยดูแลผู้ที่ได้รับผลการประเมิน 5) ความรับผิดชอบต่อสาธารณะและสวัสดิภาพของสังคม (responsibilities for general and public welfare) หมายถึง ผู้ประเมินต้องพิจารณาถึงผลประโยชน์และค่านิยมที่อาจจะมีส่วนเกี่ยวข้องกับสาธารณะและสวัสดิภาพของสังคม ปรับปรุงการปฏิบัติงานอย่างต่อเนื่อง เพิ่มความโปร่งใสภายในการประเมิน เชื่อมโยงกระบวนการประเมินกับการตัดสินใจ พิจารณาการเมืองที่เกี่ยวข้องกับการประเมิน เปิดเผยเพื่อรับข้อมูลข่าวสารจากที่อื่น สรุปได้ดังตารางที่ 2.3

ตารางที่ 2.3 สังเคราะห์มาตรฐานการประเมินที่เกี่ยวข้องกับผู้ประเมิน

Guiding Principles for Evaluators (1995)	Essential Skills Series in Evaluation (1999)	Essential Competencies for Program Evaluators (2006)
1.การดำเนินการสืบค้นอย่างเป็นระบบ (Systematic Inquiry)	การประเมินผลลัพธ์ (Evaluating for Results)	การดำเนินการประเมินอย่างเป็นระบบ (Systematic Inquiry)
2. สมรรถนะในการปฏิบัติงานของนักประเมิน (Competence)	ความเข้าใจในการประเมินโครงการ (Understanding Program Evaluation)	-สมรรถนะระหว่างบุคคล Interpersonal Competence -การปฏิบัติกับผลที่เกิดขึ้น (Reflective Practice)
3.ความซื่อสัตย์ของผู้ประเมิน (Integrity/Honesty)	การสร้างกรอบแนวทางในการประเมิน (Building an Evaluation Framework)	การจัดการโครงการ (Project Management)
4.การเคารพต่อสิทธิ ศักดิ์ศรีส่วนบุคคล (Respect for People)	การสร้างกรอบแนวทางในการประเมิน (Building an Evaluation Framework)	การฝึกปฏิบัติวิชาชีพ (Profession Practice)
5. ความรับผิดชอบต่อสาธารณะและสวัสดิภาพของสังคม (Responsibilities for General and Public Welfare)	การปรับปรุงการปฏิบัติงาน (Improving Program Performance)	การวิเคราะห์สถานการณ์ (Situational Analysis)

2.3 มาตรฐานและตัวบ่งชี้ในการประเมินอภิमानรายงานการประเมินคุณภาพภายนอกระดับอุดมศึกษา

มาตรฐานและตัวบ่งชี้ในการประเมินอภิमानรายงานการประเมินคุณภาพภายนอกระดับอุดมศึกษา ของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา

(องค์การมหาชน) (สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา องค์การมหาชน, 2550ข) ประกอบด้วย 4 มาตรฐาน 11 ตัวบ่งชี้

มาตรฐานที่ 1 ความตรงของการประเมินตามหลักเกณฑ์และวิธีการที่ สมศ. กำหนด ประกอบด้วย 2 ตัวบ่งชี้ คือ 1.1) ประเมินสอดคล้องกับความมุ่งหมายของการประเมินคุณภาพภายนอก เพื่อเป็นการตรวจสอบและยืนยันสภาพจริงในการดำเนินงานของสถาบันอุดมศึกษา ประเมินคุณภาพการศึกษาตามมาตรฐานคุณภาพ สมศ. และ 1.2) ประเมินตามระบบ หลักเกณฑ์ และวิธีการที่ สมศ. กำหนด

มาตรฐานที่ 2 ความถูกต้องสมบูรณ์ของกระบวนการประเมิน ประกอบด้วย 4 ตัวบ่งชี้ คือ 2.1) กระบวนการเก็บรวบรวมข้อมูลได้มาตรฐาน โดยมีการวิเคราะห์จากรายงานประเมินตนเองที่เป็นรายงานประจำปี เพื่อเป็นข้อมูลพื้นฐานสำหรับการประเมินคุณภาพภายนอก 2.2) ความถูกต้องครบถ้วนของข้อมูล โดยมีข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพ 2.3) ความถูกต้องของการวิเคราะห์ข้อมูล โดยมีการวิเคราะห์ข้อมูลอิงเกณฑ์มาตรฐาน อิงพัฒนาการ และข้อมูลประสิทธิผลตามแผนประจำปี/หรือแผนปฏิบัติการของสถาบันถูกต้องตามที่ สมศ.กำหนด 2.4) การแปลความหมายผลการวิเคราะห์ข้อมูลเทียบกับมาตรฐานการประเมิน

มาตรฐานที่ 3 ประสิทธิภาพของการประเมิน ประกอบด้วย 3 ตัวบ่งชี้ คือ 3.1) ความถูกต้อง เชื่อถือได้ ของผลการประเมิน 3.2) ความชัดเจนและถูกต้องของรายงานผลการประเมิน ที่มีโครงสร้างและรูปแบบตามที่ สมศ.กำหนด 3.3) ความเป็นประโยชน์และการเป็นที่ยอมรับของรายงานผลการประเมิน มีการยืนยันการยอมรับหรือไม่ยอมรับร่างรายงานผลการประเมินคุณภาพภายนอกจากสถาบัน

มาตรฐานที่ 4 ความเป็นไปได้และความเหมาะสมของข้อเสนอแนะ ประกอบด้วย 2 ตัวบ่งชี้ คือ 4.1) ความชัดเจน และความเป็นไปได้ของข้อเสนอแนะ ที่เป็นรูปธรรมและสามารถนำไปปฏิบัติได้ 4.2) ความสมเหตุสมผลของข้อเสนอแนะ ที่ตรงประเด็นปัญหาของสถาบัน สอดคล้องกับผลการประเมินคุณภาพภายนอก สอดคล้องกับอัตลักษณ์/จุดเน้นของสถาบัน และทำทนายการพัฒนาคุณภาพและมาตรฐานของสถาบัน

3. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องกับการกำหนดมาตรฐานการประเมิน แบ่งออกเป็น 2 กลุ่ม คือ 3.1) การกำหนดมาตรฐานการประเมิน จากงานวิจัยของ Laubli Loud (2004), Berends และ Roberts (2003) และ วสันต์ อดิศักดิ์ และคณะ (2549) 3.2) งานวิจัยที่นำมาตราฐานการประเมินไปประยุกต์ใช้ได้แก่ งานวิจัยของ Coryn และคณะ (2007) และ Coryn และ Scriven (2007)

3.1 การกำหนดมาตรฐานการประเมินในประเทศ สวิตเซอร์แลนด์และ ออสเตรเลีย เป็นการกำหนดมาตรฐานการประเมินให้เหมาะสมกับสภาพและบริบทการประเมินในประเทศและ นำมาตรฐานการประเมินไปประยุกต์ใช้ โดยงานวิจัยของ Laubli Loud (2004) ได้พัฒนามาตรฐานการประเมินของประเทศสวิตเซอร์แลนด์ จากการจัดตั้งสมาคมประเมินผล (Swiss Evaluation Society) เพื่อสร้างความเข้าใจในเรื่องคุณภาพของงานและการใช้ผลการประเมิน จึงพยายามพัฒนาแนวทางการประเมินและมาตรฐานในบริบทของประเทศ Swiss เรียกว่า SEVAL's Evaluation Standards โดยมี Widmer จากมหาวิทยาลัย Zurich เป็นประธานคณะกรรมการในการกำหนดมาตรฐานจากการศึกษามาตรฐานการประเมินของ The American Joint Committee on Standards for Educational Evaluation ใน 4 มาตรฐาน 28 มาตรฐานย่อย นำมาตรฐานการประเมิน มาประยุกต์ใช้จัดทำแบบตรวจสอบรายการและแบบสอบถาม เพื่อประเมินคุณภาพของการออกแบบการประเมิน และประเมินอภิปรายงานการประเมินคุณภาพในระหว่างดำเนินการ และสิ้นสุดการประเมิน ส่วนในประเทศออสเตรเลียจากงานวิจัยของ Berends และ Roberts (2003) ทำการกำหนดมาตรฐานการประเมินสำหรับประยุกต์ใช้ในโครงการชนพื้นเมืองในวิตตอเรีย ประเทศออสเตรเลีย เพื่อศึกษาผลสะท้อนของ สมาคมและมาตรฐานการประเมินออสเตรเลีย ของสมาคมการประเมินแห่งสหรัฐอเมริกา (AEA) มีจุดเน้นเพื่อประเมินโครงการเฉพาะแอลกอฮอล์และยาในชนพื้นเมือง การประเมินนี้ให้ใช้มาตรฐานอรรถประโยชน์ของ AEA หากผู้มีส่วนได้ส่วนเสียหลักในการเข้าร่วมการประเมิน พิจารณาความสอดคล้องของวิธีกับมาตรฐานอรรถประโยชน์และเสนอแนะข้อบ่งชี้ที่จะปรับปรุงมาตรฐานที่จะเป็นประโยชน์ ตามแนวทางของสมาคมประเมินของแอฟริกา (AEG) ผลสะท้อนจากการประเมิน ทำให้ระบุข้อบ่งชี้เพื่อเปลี่ยนแปลงโครงการชนพื้นเมืองในอนาคต

สำหรับงานวิจัยที่เกี่ยวข้องกับวิธีการพัฒนามาตรฐานแห่งชาติ มีขั้นตอนในการพัฒนามาตรฐานเพื่อให้มีความน่าเชื่อถือ เหมาะสมกับสภาพการประเมินในข้อบ่งชี้ที่ทำการประเมิน ซึ่ง วสันต์ อดิศักดิ์ และคณะ (2549) ทำการพัฒนามาตรฐานแห่งชาติทางเทคโนโลยีการศึกษาสำหรับสถาบันผลิตบัณฑิตทางการศึกษา โดยมีขั้นตอนการวิจัย 5 ขั้นตอน คือ 1) การกำหนดกรอบสำหรับการพัฒนามาตรฐานทางเทคโนโลยีการศึกษา 2) การพัฒนาร่างมาตรฐานทางเทคโนโลยีการศึกษา 3) การศึกษาความคิดเห็นผู้บริหารสำหรับผลิตบัณฑิตทางการศึกษา 4) การวิพากษ์มาตรฐานทางเทคโนโลยีการศึกษา 5) การรับรองมาตรฐานทางเทคโนโลยีการศึกษาโดยสภาคุณบดีคณะครุศาสตร์ศึกษาแห่งประเทศไทย ผลการวิจัยได้เสนอมาตรฐานแห่งชาติทางเทคโนโลยีการศึกษา สำหรับสถาบันผลิตบัณฑิตทางการศึกษา 9 มาตรฐาน 34 ตัวบ่งชี้ ดังนี้ 1) มาตรฐานด้าน

ภาวะผู้นำทางเทคโนโลยีการศึกษาของสถาบันผลิตบัณฑิตทางการศึกษา 2) มาตรฐานด้านโครงสร้างของหลักสูตรการผลิตบัณฑิตทางการศึกษา 3) มาตรฐานด้านโครงสร้างพื้นฐานทางเทคโนโลยีทางการศึกษา 4) มาตรฐานด้านเทคโนโลยีร่วมสมัยสนับสนุนการเรียนการสอน 5) มาตรฐานด้านบุคลากรทางเทคโนโลยีการศึกษา 6) มาตรฐานด้านสมรรถนะทางเทคโนโลยีการศึกษาของอาจารย์ 7) มาตรฐานด้านรูปแบบการจัดการเรียนการสอนที่เอื้อและส่งเสริมการพัฒนาสมรรถนะทางการใช้เทคโนโลยีการศึกษาแก่นิสิต นักศึกษาครู 8) มาตรฐานด้านการฝึกประสบการณ์วิชาชีพ และ 9) มาตรฐานด้านสมรรถนะทางเทคโนโลยีการศึกษาของบัณฑิตทางการศึกษา

3.2 งานวิจัยเกี่ยวกับการนำมาตรฐานการประเมินไปประยุกต์ใช้พัฒนาโมเดลและประเมินโมเดลตามมาตรฐานการประเมิน คือ งานวิจัยของ Coryn และคณะ (2007) ทำการศึกษาเกี่ยวกับโมเดลและกลไกสำหรับประเมินทุนวิจัยของรัฐบาล เปรียบเทียบระหว่างประเทศ เพื่อบรรยายจำแนก และประเมินเปรียบเทียบ โมเดลและกลไกระดับชาติที่ใช้ประเมินวิจัยและจัดสรรทุนวิจัยใน 16 ประเทศ คือ ออสเตรเลีย เบลเยียม เชโกสโลวาเกีย ฟินแลนด์ ฝรั่งเศส เยอรมัน ฮังการี ไอร์แลนด์ ญี่ปุ่น เนเธอร์แลนด์ นิวซีแลนด์ โปแลนด์ สวีเดน สหราชอาณาจักรอังกฤษ และสหรัฐอเมริกา จำแนกออกเป็น 5 โมเดล มีการออกแบบการวิจัย เป็นวิธีการวิจัยและประเมินโดยใช้การกำหนดมาตรฐานบุคคลากร และเคลฟายเทคนิคเพื่อปรับปรุงการศึกษาวิจัยที่เหมาะสมใน 2 ขั้นตอน คือ ขั้นแรก แนะนำผู้ทำการตัดสินใจและกระบวนการให้คะแนนอย่างอิสระ และขั้นที่สอง ออกแบบเพื่อค้นหาประสามติในการให้คะแนนตัดสินใจ โดยใช้กลุ่มผู้ตัดสินใจต่างกันซึ่งยังไม่เคยเข้ามีส่วนร่วมในการศึกษา การวัดคุณลักษณะทั่วไปร่วมกันอย่างใกล้ชิดที่สัมพันธ์กับทุนเพื่อวัดการปฏิบัติงานที่ผ่านมา โมเดลและกลไกของประเทศทั้ง 16 ประเทศถูกให้คะแนนอย่างอิสระ ประกอบด้วย 25 ตัวบ่งชี้ การให้ระดับคะแนนใน 5 มาตรฐาน มาตรฐานละ 5 ตัวบ่งชี้ ได้แก่ มาตรฐานความตรง (validity) มาตรฐานความน่าเชื่อถือของผู้ประเมิน (credibility) มาตรฐานอรรถประโยชน์ (utility) มาตรฐานประสิทธิผลต้นทุน (cost-effectiveness) และมาตรฐานจริยธรรมในการประเมิน (ethicality) ระดับคะแนนสูงสุดได้จากโมเดล มาตรฐานวัดประเมินงานวิจัย ในขณะที่โมเดล Bulk funding และ indicator-driven ได้ระดับคะแนนต่ำสุด ซึ่งหมายความว่า การปฏิบัติและนโยบายการประเมินงานวิจัยจะได้รับการพิจารณาและวิพากษ์

นอกจากนี้ Coryn และ Scriven (2007) ทำการศึกษาวิจัยเพิ่มเติม โดยเปรียบเทียบโมเดลประเมินการวิจัยในระดับชาติต่างๆ ใน 5 มาตรฐานดังกล่าวข้างต้น 25 ตัวบ่งชี้คุณภาพ โดยใช้ นักวิจัยอาชีพและนักประเมินใน 2 ประเทศ (สหรัฐอเมริกา และนิวซีแลนด์) เป็นผู้ตัดสินใจ พบว่า มาตรฐานด้านความตรง (validity) มีความแตกต่างของคะแนนความตรงระหว่างนักประเมินเป็น

0.81% สัมประสิทธิ์สหสัมพันธ์ระหว่างนักประเมินทำการประเมินในประเทศเดียวกันเท่ากับ .90 สัมประสิทธิ์สหสัมพันธ์ภายในทีมเท่ากับ .90 มาตรฐานด้านความน่าเชื่อถือ (credibility) มีความแตกต่างของคะแนนความน่าเชื่อถือระหว่างนักประเมินเป็น 1.18% สัมประสิทธิ์สหสัมพันธ์ระหว่างนักประเมินทำการประเมินในประเทศเดียวกันเท่ากับ .92 สัมประสิทธิ์สหสัมพันธ์ภายในทีมเท่ากับ .91 มาตรฐานด้านการนำไปใช้ประโยชน์ (utility) มีความแตกต่างของคะแนนการนำไปใช้ประโยชน์ระหว่างนักประเมินเป็น 0.37% สัมประสิทธิ์สหสัมพันธ์ระหว่างนักประเมินทำการประเมินในประเทศเดียวกันเท่ากับ .93 สัมประสิทธิ์สหสัมพันธ์ภายในทีมเท่ากับ .93 มาตรฐานด้านประสิทธิภาพ (cost-effectiveness) มีความแตกต่างของคะแนนประสิทธิภาพระหว่างนักประเมินเป็น 1.37% สัมประสิทธิ์สหสัมพันธ์ระหว่างนักประเมินทำการประเมินในประเทศเดียวกันเท่ากับ .84 สัมประสิทธิ์สหสัมพันธ์ภายในทีมเท่ากับ .83 และมาตรฐานด้านจริยธรรม (ethicality) มีความแตกต่างของคะแนนจริยธรรมระหว่างนักประเมินเป็น 2.62 % สัมประสิทธิ์สหสัมพันธ์ระหว่างนักประเมินทำการประเมินในประเทศเดียวกันเท่ากับ .78 สัมประสิทธิ์สหสัมพันธ์ภายในทีมเท่ากับ .76 เมื่อพิจารณาเปรียบเทียบผลการประเมินระหว่างประเทศ พบว่า ประเทศเนเธอร์แลนด์ นิวซีแลนด์ สหราชอาณาจักรอังกฤษ และสหรัฐอเมริกาส่วนมากมีค่าคะแนนสูงทุกมาตรฐาน และประเทศฮ่องกงมีค่าคะแนนสูงในมาตรฐานความน่าเชื่อถือ และการนำไปใช้ประโยชน์ ส่วนประเทศฝรั่งเศสมีค่าคะแนนต่ำสุดในทุกมาตรฐาน

ตอนที่ 3 แนวคิดทฤษฎีเกี่ยวกับการประเมินอภิमान

แนวคิดทฤษฎีเกี่ยวกับการประเมินอภิमान ในตอนนี้แบ่งออกเป็น 6 หัวข้อ ได้แก่ 1) ความหมายและวัตถุประสงค์ของการประเมินอภิमान 2) หลักการประเมินอภิमान 3) ประเภทของการประเมินอภิमान 4) วิธีการที่ใช้ในการประเมินอภิमान 5) การสังเคราะห์มาตรฐานการประเมินอภิमान และ 6) งานวิจัยที่เกี่ยวข้อง

1. ความหมายและวัตถุประสงค์ของ การประเมินอภิमान

การประเมินอภิमानมาจากคำศัพท์ในภาษาอังกฤษว่า meta-evaluation และ evaluation of evaluations และ second order evaluation ที่บัญญัติขึ้นโดย Scriven (1969) มาจากแนวคิดของ Glass ตามคำศัพท์ “meta-analysis” ที่หมายถึงการวิจัยงานวิจัย จากแนวคิดของ Scriven (1969) ได้นิยาม การประเมินอภิमानว่า เป็นการประเมินงานประเมิน ระบบการประเมิน หรือ ความสามารถในการประเมิน ซึ่ง Stufflebeam (2000) นิยาม การประเมินอภิमानว่า เป็นกระบวนการรวบรวมข้อมูล การประยุกต์ใช้สารสนเทศเชิงบรรยาย และสารสนเทศเชิงตัดสินใจ

เกี่ยวกับอรรถประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้องของการประเมินเพื่อเป็นแนวทางในการประเมินและเผยแพร่รายงานจุดแข็ง จุดอ่อน ผู้สาคาธาระณะ สำหรับคำว่า “การประเมินอภิมาน” ศาสตราจารย์ ดร.สมหวัง พิธิยานุวัฒน์ ได้บัญญัติศัพท์ขึ้น จากแนวคิดเดียวกับการบัญญัติศัพท์ “การวิเคราะห์ อภิมาน” ที่บัญญัติขึ้นโดยคณะกรรมการบริหารหลักสูตรปริญญาตรีบัณฑิต สาขาพัฒนศึกษาศาสตร์และสาขาการวิจัยและพัฒนาหลักสูตร มหาวิทยาลัยศรีนครินทรวิโรฒ โดยสมหวัง พิธิยานุวัฒน์ (2549) ให้ความหมายของการประเมินอภิมานว่า การตัดสินคุณค่าภายนอกและคุณค่าภายในของงานประเมิน ศิริชัย กาญจนวาสี (2547ก) ให้ความหมายว่า เป็นการตัดสินคุณภาพของการประเมิน โดยคุณภาพประกอบด้วย การกำหนดคำถามของการประเมิน วิธีการประเมิน และผลการประเมิน ซึ่งการประเมินอภิมานอาจเป็นกลไกของการควบคุมคุณภาพของการประเมิน สามารถกระทำได้ด้วยการติดตามตรวจสอบการประเมิน การวิพากษ์วิจารณ์ โดยเปรียบเทียบกับเกณฑ์มาตรฐานหรือทำการประเมินซ้ำ นางลักษณ วิรัชชัย (2550) ได้สรุปความหมายของการประเมินอภิมานจากแนวคิดของนักวิชาการหลายท่าน ดังนี้ การประเมินอภิมาน หมายถึง กระบวนการที่มีระบบประกอบด้วย การกำหนดข้อมูลสารสนเทศที่ต้องการ การรวบรวมข้อมูลสารสนเทศจากผลการประเมินงานประเมินหรือนักประเมิน เพื่ออธิบาย ให้คุณค่า ตัดสินคุณค่า ความเหมาะสมสอดคล้องของงานประเมินหรือนักประเมินกับเกณฑ์และมาตรฐานที่กำหนดไว้ วิเคราะห์หาจุดอ่อน จุดแข็งในการดำเนินงาน รวมทั้งสังเคราะห์ผลการประเมินงานประเมินหรือนักประเมิน ทั้งในส่วนผลการประเมินเดิมและการประเมินอภิมาน

การประเมินอภิมานมีวัตถุประสงค์สำคัญของ 2 ประการ คือ ประการแรก เพื่อตรวจสอบคุณภาพ การควบคุมกำกับ และการรับรองคุณภาพของโครงการประเมินทั้งโครงการเดียวหรือหลายโครงการ หรือเป็นการประเมินศักยภาพของนักประเมินเพียงคนเดียว หรือหลายคน ประการที่สอง เพื่อสังเคราะห์สรุปผลการประเมินกรณีที่มีการประเมินหลายโครงการหรือมีผลการประเมินจากนักประเมินหลายคน ให้ได้ผลการประเมินนำไปใช้ประโยชน์ต่อไป ใช้ในการพัฒนาคุณภาพการประเมิน ทั้งของนักประเมินและคุณภาพของหน่วยงาน Migotsky และ Stake (2001) เสนอให้เห็นว่าการประเมินอภิมานมีวัตถุประสงค์ที่สำคัญ 4 ประการ คือ ประการแรก เพื่อประเมินระดับของการบรรลุเป้าหมายที่กำหนดไว้ของโครงการประเมิน ประการที่สอง เพื่อประเมินผลกระทบของโครงการประเมินต่อบุคคลหรือกลุ่มที่เป็นเป้าหมายของการประเมิน ประการที่สาม เพื่อประเมินประสิทธิภาพและประสิทธิผลของโครงการประเมิน ประการที่สี่ เพื่อนำเสนอข้อมูลสารสนเทศสำหรับการปรับปรุงโครงการประเมินให้ดียิ่งขึ้น ส่วน Scriven (2003) และ Stufflebeam (2002) ได้สรุปวัตถุประสงค์ของการประเมินอภิมาน คือ เพื่อตรวจสอบ ควบคุม และรับรองคุณภาพของการประเมิน

หลายโครงการหรือนักประเมินหลายคน รวมถึงเพื่อสังเคราะห์สรุปผลการประเมินกรณีที่มีการประเมินหลายโครงการ **ดังนั้น** สรุปได้ว่า การประเมินอภิมานมีวัตถุประสงค์หลัก คือ เพื่อควบคุม ตรวจสอบ และรับรองคุณภาพงานประเมินหรือนักประเมิน และเพื่อสังเคราะห์ผลการประเมินจากหลายๆโครงการประเมิน โดยทำการประเมินการบรรลุเป้าหมาย ผลกระทบ ประสิทธิภาพและประสิทธิผลเพื่อนำผลการประเมินไปปรับปรุงพัฒนาต่อไป

ความจำเป็นที่จะต้องทำการประเมินอภิมาน Stufflebeam (2001) เสนอว่า เนื่องจากศาสตร์ทางการประเมินมีความก้าวหน้าในวิธีวิทยาและการบริการสาธารณะ โดยนักประเมินสามารถทำการประเมินอภิมานอย่างเป็นระบบได้ การประเมินอภิมานเป็นกระบวนการวิเคราะห์ การค้นพบ และการประยุกต์เพื่อบรรยายสารสนเทศ และการตัดสินสารสนเทศเกี่ยวกับมาตรฐานการประเมินใน ด้านอรรถประโยชน์ ความเป็นไปได้ ความเหมาะสม และความถูกต้อง และธรรมชาติระบบ สมรรถนะ ความซื่อสัตย์ การยกย่องนับถือ และความรับผิดชอบต่อสังคม เพื่อเป็นแนวทางในการประเมินและรายงานจุดแข็ง จุดอ่อนสู่สาธารณะ การประเมินอภิมาน ความก้าวหน้า (formative metaevaluations) เป็นการยอมรับและการทำการประเมิน เพื่อช่วยนักประเมินวางแผน ดำเนินการ ปรับปรุง แปลความและรายงานการศึกษาประเมิน ส่วนการประเมินอภิมานผลสรุป (summative metaevaluations) ดำเนินการติดตามการประเมินเพื่อช่วยผู้รับผลการประเมินเห็นจุดแข็งและจุดอ่อนของการประเมิน และตัดสินคุณค่าภายในและภายนอก การประเมินอภิมานเป็นที่เปิดเผย ได้รับการยอมรับและความสนใจของสถาบัน เพื่อประกันว่าการประเมินให้ข้อค้นพบและข้อสรุป ซึ่งการทำการประเมินอย่างต่อเนื่องเพื่อปรับปรุง และการบริหารสถาบัน ให้ระบบการประเมินอย่างมีประสิทธิภาพและประสิทธิผล นักประเมินมืออาชีพมีหน้าที่ทำการประเมินอภิมานอย่างเข้มงวดมากยิ่งขึ้น แต่จำเป็นต้องเพิ่มเครื่องมือและกระบวนการที่จะประยุกต์ใช้มาตรฐานและหลักการของแนวปฏิบัติการประเมินที่ดี

2. หลักการประเมินอภิมาน

หลักการประเมินอภิมานตามที่ Stufflebeam (1974) สรุปไว้ มีรายละเอียดดังต่อไปนี้ คือ การประเมินอภิมานเป็นการประเมินคุณค่าภายใน (merit) และคุณค่าภายนอก (worth) ของงานประเมินหรือสิ่งที่ถูกประเมิน เพื่อใช้ประโยชน์ต่อการตัดสินใจ ความรับผิดชอบที่ตรวจสอบได้ และการประกันการทำงาน รวมทั้งประโยชน์ในการใช้ผลการประเมินเพื่อการปรับปรุงงาน ดังนั้นคุณค่าในการประเมินอภิมานจึงมีประโยชน์ในการสนับสนุนการตัดสินใจเชิงรุกเกี่ยวกับงานประเมิน และให้สารสนเทศเพื่อปรับปรุงงานประเมินในแต่ละขั้นตอน ซึ่งการประเมินอภิมาน เป็นการประเมินวัตถุประสงค์ การออกแบบหรือการวางแผนการดำเนินงาน วิธีดำเนินงานและผลการดำเนินงาน

ซึ่งสามารถประเมินความก้าวหน้าและประเมินภาพรวม ความสำคัญของการประเมินอภิमान เพื่อประเมินวัตถุประสงค์การประเมิน ความเหมาะสมของการออกแบบการประเมินและการวางแผนการประเมิน ความพอเพียงในการดำเนินการประเมิน ความสำคัญ และคุณภาพของผลการประเมิน ดังนั้นการประเมินอภิमानให้สารสนเทศ สำหรับการประเมินโดยทั่วไปจะให้สารสนเทศเกี่ยวกับลักษณะของโครงการ สารสนเทศในการตัดสินคุณภาพของโครงการ ข้อเสนอแนะเพื่อใช้ในการปรับปรุงโครงการ ดังนั้น ในการประเมินอภิमानจึงให้สารสนเทศที่บรรยายลักษณะของการประเมิน ผลการประเมินงานประเมิน ข้อเสนอแนะเพื่อใช้ในการพัฒนาปรับปรุงวิธีการประเมินและการใช้ประโยชน์จากผลการประเมิน โดยที่ผู้ใช้ผลการประเมินอภิमानเป็นกลุ่มเช่นเดียวกับผู้ใช้ผลการประเมินทั่วไป คือ ผู้ที่เกี่ยวข้อง ผู้มีส่วนได้ส่วนเสีย โดยรวมถึง นักประเมินหรือผู้สนใจการประเมินทุกกลุ่ม ทุกระดับด้วย ดังนั้นบุคคลที่ควรเป็นนักประเมินอภิमानควรมาจากทั้งนักประเมินภายใน และนักประเมินภายนอก เพื่อทำการประเมินอภิमान โดยการประเมินกลุ่มแรกเรียกว่า การประเมินอภิमानงานประเมินภายใน กระทำโดยกลุ่มนักประเมินภายในองค์กร หรือสถานศึกษา และการประเมินกลุ่มที่สอง เรียกว่า การประเมินอภิमानงานประเมินภายนอก อาจดำเนินการโดยกลุ่มนักประเมินภายนอก หรือกลุ่มนักประเมินภายในที่มาจากหน่วยงานต้นสังกัดแต่อยู่ภายนอกองค์กรหรือสถานศึกษา โดยมีวิธีการประเมินอภิमानประกอบด้วยขั้นตอนที่สำคัญ 3 ขั้นตอน คือ 1) การกำหนดวัตถุประสงค์และคำถามสำหรับการประเมินอภิमान 2) การรวบรวมและวิเคราะห์ข้อมูลเพื่อตอบคำถามการประเมินอภิमानโดยการตัดสินคุณค่างานประเมินตามเกณฑ์การประเมินอภิमानที่กำหนด 3) การนำผลการประเมินอภิमानไปใช้ประโยชน์

โดยที่ประเมินอภิमान ใช้มาตรฐานในการประเมินด้านอรรถประโยชน์ (Utility) ความเป็นไปได้ (Feasibility) ความเหมาะสม (Propriety) และความถูกต้อง (Accuracy) รวมถึงประสิทธิภาพและประสิทธิผลของการประเมิน และยังใช้มาตรฐานของ The Joint Committee on Standards for Educational Evaluation อีกด้วย ดังที่ Stufflebeam (1999) ได้พัฒนาแบบตรวจสอบรายการสำหรับใช้ในการประเมินอภิमानตามมาตรฐานของ the Joint committee on standards for evaluation มีรายละเอียด 4 ด้าน คือ 1) มาตรฐานด้านอรรถประโยชน์ (Utility Standards) 7 มาตรฐานย่อย 2) มาตรฐานด้านความเป็นไปได้ (Feasibility Standards) 3 มาตรฐานย่อย 3) มาตรฐานด้านความเหมาะสม (Propriety Standards) 8 มาตรฐานย่อย 4) มาตรฐานด้านความถูกต้อง (Accuracy Standards) 12 มาตรฐานย่อย รายละเอียดในแต่ละมาตรฐานได้นำมากำหนดเป็นมาตรฐานการประเมินสำหรับการประเมินงานประเมินโครงการ (Program Evaluation Meta-evaluation Checklist) รวม 30 มาตรฐาน

ศิริชัย กาญจนวาสี (2547ก) เสนอว่า การประเมินอภิमानควรกำหนดเกณฑ์ ตัดสินคุณภาพการประเมินให้ครอบคลุมการประเมินคุณภาพ 3 ด้าน คือ คำถามการประเมิน วิธีการประเมิน และผลการประเมิน ซึ่ง**การกำหนดคำถามการประเมิน** ที่ตรงประเด็นกับความ ต้องการจำเป็นที่แท้จริงของการจัดกิจกรรมหรือโครงการนั้นๆ เพื่อใช้เป็นแนวทางสู่การตัดสิน คุณค่าของสิ่งที่ต้องการให้เกิดขึ้นที่แท้จริง ตามเป้าหมายของการจัดกิจกรรมหรือโครงการ ใน ขณะเดียวกัน คำถามการประเมินควรสนองความต้องการใช้สารสนเทศของผู้ที่เกี่ยวข้องสำหรับ การตัดสินใจปรับปรุง/เปลี่ยนแปลงการจัดกิจกรรมหรือโครงการเพื่อพัฒนาสังคมต่อไปในอนาคต ดังนั้นการวิเคราะห์ถึงความต้องการจำเป็นที่แท้จริง และการมีส่วนร่วมในการกำหนดคำถามการ ประเมินของผู้ต้องการใช้สารสนเทศ จึงมีความสำคัญต่อคุณภาพของคำถามการประเมิน **วิธีการ ประเมิน** ควรเป็นกระบวนการที่ตอบสนองต่อจุดมุ่งหมายของการประเมิน สามารถจัดกระทำให้ บรรลุผลของการตัดสินคุณค่าของสิ่งที่มุ่งประเมินได้อย่างถูกต้องหรือเหมาะสม เนื่องจากคุณค่า ของสิ่งที่มุ่งประเมินนั้นขึ้นอยู่กับบริบท วิธีการประเมินจึงควรมีลักษณะที่สอดคล้องและเหมาะสม กับบริบท ตลอดจนเป็นที่เข้าใจและยอมรับได้ของผู้ที่เกี่ยวข้อง ดังนั้นจะเห็นได้ว่าองค์ประกอบ สำคัญที่จะนำไปสู่การเลือกวิธีการประเมินที่เหมาะสม ได้แก่ การกำหนดวัตถุประสงค์และเป้าหมาย ของการประเมินที่ชัดเจน สภาพแวดล้อมของการประเมินประสบการณ์/ ค่านิยมของนักประเมินและ ผู้เกี่ยวข้อง **ผลการประเมิน** ควรมีคุณลักษณะสำคัญที่สุด คือ สามารถนำไปใช้ประโยชน์ได้ เป็นประโยชน์ต่อผู้เกี่ยวข้องอย่างกว้างขวาง นักประเมินมีหน้าที่ที่จะต้องสามารถระบุตั้งแต่เริ่มต้นทำ การประเมินว่าใครบ้างที่อยู่ในข่ายจะเป็นผู้ใช้ผลการประเมิน เพื่อนักประเมินจะได้สร้างสัมพันธภาพ ที่ดีกับบุคคลเหล่านั้น จะได้ทราบว่าเขาต้องการข้อมูลอะไรบ้าง และเป็นแนวทางสู่ความร่วมมือ ด้านต่างๆ การให้ความสนใจ การยอมรับผลการประเมิน และการใช้ผลการประเมินให้เป็น ประโยชน์ ซึ่งผลการประเมินควรมี**ความยุติธรรม** ในการตัดสินคุณค่าของสิ่งที่มุ่งประเมินโดย จะต้องมีการเก็บรวบรวมข้อมูลอย่างรอบด้าน และจากผู้เกี่ยวข้องหลายๆฝ่าย มีการใช้เกณฑ์การ ตัดสินที่สมเหตุสมผลและยอมรับได้ พร้อมทั้งมีการเชื่อมโยงคุณค่าของสิ่งที่ประเมินภายใต้บริบท นั้น สู่คุณค่าของสิ่งนั้นในสายตาของสาธารณชน นอกจากนี้ผลการประเมินควรมี**ความน่าเชื่อถือ** คือ สามารถเปิดเผยความจริงเกี่ยวกับคุณค่าของสิ่งที่ประเมินได้ ตัวบ่งชี้ที่สำคัญที่บ่งบอกถึงความ น่าเชื่อถือของผลการประเมินอาจพิจารณาได้จากความน่าเชื่อถือของ “**นักประเมิน**” ในแง่ความ เชื่อวาทะในระเบียบวิธีการประเมิน เนื้อหา และจรรยาบรรณ ความสมเหตุสมผลของการกำหนด วัตถุประสงค์ของการประเมิน มีการออกแบบและใช้วิธีการประเมินที่เชื่อถือได้

นอกจากนี้ Davidson (2005) ได้เสนอว่าการประเมินอภิमानควรประกอบด้วยมิติคุณค่าหลัก 5 ด้าน ได้แก่ 1) ความตรงของการประเมิน (validity) 2) อรรถประโยชน์ (utility) 3) จริยธรรมในการประเมิน (ethicality) 4) ความน่าเชื่อถือของผู้ประเมิน (credibility) และ 5) ประสิทธิภาพต้นทุน (cost-effectiveness) ซึ่งรายงานการประเมินจะเป็นการสรุปสารสนเทศที่ได้จากการประเมินตามเกณฑ์ที่ทำให้เกิดความยุติธรรมในการประเมินอภิमानจึงประกอบด้วย 5 มาตรฐานดังกล่าว

ความตรงของการประเมิน (validity) เป็นการตัดสินใจข้อสรุปผลการประเมินอย่างถูกต้องสมเหตุสมผล ซึ่งการประเมินอภิमानเป็นการตรวจสอบข้อสรุปจากทีมประเมินทั้งข้อมูลเชิงบรรยายและคุณค่าที่สอดคล้องกับเกณฑ์มากที่สุด เกณฑ์ในการประเมินที่สำคัญพิจารณาจากความสำคัญมากที่สุด และการปฏิบัติงานพิจารณาจากความพึงพอใจ ดีหรือดีมาก ซึ่งการประเมินอภิमानมีความสำคัญที่จะตรวจสอบแหล่งที่ประเมินอย่างละเอียด และตรวจสอบข้อมูลทั้งเชิงบรรยายและคุณค่าตามเกณฑ์ โดยทีมประเมินทุกทีมจะปฏิบัติงานภายใต้ข้อกำหนดเกี่ยวกับการบริหารจัดการ ทุน/งบประมาณ กำหนดเวลา และอื่นๆ รวมถึงวิธีการในการประเมิน สิ่งเหล่านี้ต้องนำไปพิจารณาในการประเมินอภิमान เครื่องมือที่ใช้ในการตรวจสอบความตรงในการประเมิน ได้แก่ แบบตรวจสอบรายการของ Scriven (2003) ซึ่งช่วยตัดสินใจที่ครอบคลุมรายการการประเมิน มีความสำคัญในการประเมินคุณภาพการประเมิน นอกจากนี้ยังมีเครื่องมืออื่นที่เป็นที่รู้จักเพื่อใช้ในการพิจารณาความตรงการประเมินให้เป็นที่น่าเชื่อถือ ประเมินสิ่งที่มุ่งประเมิน คือ Program Evaluation Standards, Personnel Evaluation Standards และ Student Evaluation Standards ที่เผยแพร่โดย the Joint Committee on Standards for Educational Evaluation ซึ่งมาตรฐานดังกล่าวครอบคลุมมาตรฐาน 4 ด้าน คือ อรรถประโยชน์ (utility) ความเหมาะสม (propriety) ความเป็นไปได้ (feasibility) และความถูกต้อง (accuracy) ซึ่งในมาตรฐานด้านความถูกต้องจะกล่าวถึงความตรงของการประเมินและการใช้ประโยชน์จากการพิจารณารายการในการประเมินหลายรายการ สำหรับตัวอย่างในการพิจารณาแบบตรวจสอบรายการกับ Joint Committee Standards มีเกณฑ์ที่ใช้พิจารณาประเมินความตรงของการประเมิน ได้แก่ ประเมินครอบคลุมทุกแหล่ง ได้แก่ ความต้องการจำเป็น มาตรฐานที่ตรงประเด็น นิยามคุณค่าโดยทั่วไป สรุปครอบคลุมทั้งกระบวนการ ผลลัพธ์และความคุ้มค่าของการประเมิน ข้อสรุปที่ไม่ตรงประเด็นหรือคลาดเคลื่อนจากเกณฑ์ ข้อมูลที่ใช้โดยตรงกับเกณฑ์ ข้อสรุปในการวิเคราะห์ทั้งเชิงคุณภาพและเชิงปริมาณ ที่ได้จากข้อมูลที่ต้องการ ข้อมูลที่แปลความได้อย่างชัดเจน ความชัดเจนเกี่ยวกับข้อสรุปในการประเมินที่มาจากข้อมูลที่ไม่เป็นตรรกะ ข้อสรุปที่สนับสนุนสมเหตุสมผล

อรรถประโยชน์ (utility) การประเมินที่ดีจะเป็นประโยชน์ต่อผู้รับผลการประเมิน ซึ่งอย่างน้อยข้อค้นพบจะต้อง 1) ตรงกับคำถามหรือสอดคล้องกับความต้องการของผู้รับผลการประเมิน 2) ข้อค้นพบที่ได้ตรงกับช่วงเวลาที่ใช้ผลต้องการใช้ในการตัดสินใจ 3) ใช้การสื่อสารที่ชัดเจนด้วยภาษา สื่อและช่องทางการสื่อสาร โดยการประเมินแบบมีส่วนร่วมสามารถได้รับประโยชน์จากกระบวนการประเมิน ภายใต้อัตลักษณ์การประเมินใหม่ๆ หรือความรู้สำหรับบุคคลที่มีส่วนร่วมในการทำการประเมิน การเพิ่มความสามารถในการเรียนรู้ทำให้เห็นผลลัพธ์ของการประเมินด้วยตนเอง หรือบางครั้งการประเมินอาจมีประโยชน์ต่อองค์กร ผู้มีส่วนได้ส่วนเสีย และ/หรือ ชุมชน ส่วนที่สำคัญของประโยชน์จากการประเมินจากข้อค้นพบที่ได้ ข้อวินิจฉัยที่ดีจากการประเมินอภิमानสามารถช่วยแสดงเหตุผลที่เกี่ยวข้องจุดเน้นหรือทำการประเมินหรือสื่อสารข้อค้นพบ สิ่งเหล่านี้จะช่วยให้ทีมประเมินปรับปรุงการใช้ประโยชน์สำหรับการประเมินในครั้งต่อไป นักทฤษฎีและนักปฏิบัติหลายคนได้แย้งถึงเกณฑ์ที่ใช้ในการประเมินอภิमानที่ไม่ได้พิจารณาจากการควบคุมผู้ประเมิน ในความเป็นจริงแล้วผลลัพธ์ทั้งหมดจากการประเมินเป็นผลมาจากการควบคุมจากบุคคล โปรแกรม หรือทีมที่ทำการประเมิน

จริยธรรมในการประเมิน (ethicality) เป็นการประเมินที่สมบูรณและยุติธรรมต่อผู้รับผลการประเมิน การประเมินเป็นหนทางซึ่งปฏิบัติในส่วนของกระบวนการสำหรับการประเมินอภิमान ซึ่งคุณค่าของการประเมินที่สอดคล้องกับเกณฑ์นั้นประยุกต์ใช้ในส่วนของกฎหมาย จรรยาบรรณการประเมินและมาตรฐานทางอาชีพ ความเหมาะสมและความลึกลับทางวัฒนธรรม การประยุกต์ใช้มาตรฐานทางกฎหมายเพื่อให้เกิดความมั่นใจได้ว่า การประเมินสอดคล้องกับเกณฑ์ของชาติ นานาชาติ และกฎหมายท้องถิ่น ส่วนมาตรฐานทางจริยธรรม ได้แก่ ไม่เป็นอันตรายต่อผู้มีส่วนร่วม เป็นประโยชน์ต่อผู้รับผลการประเมิน และการปกปิดความลับของผู้ถูกประเมิน มาตรฐานในการประเมินวิชาชีพรวมถึง The American Evaluation Association's *Guiding Principles for Evaluators* ของสมาคมประเมินผลแห่งสหรัฐอเมริกา และสอดคล้องกับมาตรฐานการประเมินของ The Joint Committee on Standard for Educational Evaluation นอกจากนี้ความเหมาะสมทางวัฒนธรรม อาจประเมินโดยใช้แนวทางที่สอดคล้องกับบริบทของวัฒนธรรมนั้นๆ อันได้แก่ การนำชุมชนเข้ามามีส่วนร่วมในการประเมินซึ่งเป็นแนวคิดที่ดี เพื่อสร้างความมั่นใจให้แก่ทีมผู้ประเมินให้สามารถเข้าใจถึงบริบทในการประเมินได้อย่างดี แต่ถ้าหากช่วงเวลาในการประเมินกับช่วงเวลาชุมชนที่จะเข้าร่วมประเมินได้ไม่สอดคล้องกัน อาจให้ชุมชนเป็นที่ปรึกษา เพื่อจัดหาข้อมูลในการประเมินในหลายๆขั้นตอนในการทำการประเมิน

ความน่าเชื่อถือของนักประเมิน (credibility) เป็นการทำการประเมินใช้เพื่อประเมินความรับผิดชอบที่ตรวจสอบได้ หรือการเรียนรู้ภายใน เพื่อขยายข้อค้นพบที่ได้รับจากแหล่งข้อมูล ความน่าเชื่อถือของนักประเมินขึ้นอยู่กับ 3 ส่วนที่สำคัญ ดังนี้ 1) ความคุ้นเคยกับบริบท รวมถึงความรู้เกี่ยวกับองค์กรหรือโครงการประเมิน (ได้แก่ การเมือง วัฒนธรรมองค์กร โครงสร้างองค์กร เป็นต้น) หรือ ชุมชน หรือ สังคม ที่สนับสนุนสิ่งที่มุ่งประเมิน (เช่น ประวัติศาสตร์ วัฒนธรรม สาธารณูปโภค การเข้าถึงทรัพยากร) ความคุ้นเคยกับบริบทที่มีความสำคัญ เมื่อรายงานการประเมินต่อผู้รับผลการประเมินภายใน และชุมชน 2) ความอิสระ ความยุติธรรม และ/หรือ ไม่มีผลประโยชน์ทับซ้อน มีแนวโน้มที่มีความสำคัญมาก ถ้าทีมการประเมิน รายงานผลการประเมินต่อผู้รับผลการประเมินภายใน หรือผู้มีส่วนได้ส่วนเสียที่สนใจหรือมีมุมมองขัดแย้ง กับกรรมการประเมิน แม้ว่าจะเป็นการเฉพาะ หรือมีผลประโยชน์ทับซ้อนที่สามารถประนีประนอมได้ ความน่าเชื่อถือของข้อค้นพบอาจถูกนำไปใช้เพียงเล็กน้อย และ 3) ความเชี่ยวชาญในการประเมินและการสืบค้นเรื่องราวต่างๆ มีแนวโน้มที่จะมีความสำคัญสำหรับความน่าเชื่อถือในการกำหนดบางอย่างมากกว่าสิ่งอื่นๆ โดยทั่วไปผู้รับผลการประเมินมีแนวโน้มที่สนใจความเชี่ยวชาญในเนื้อหา มากกว่าความเชี่ยวชาญทางการประเมิน เพราะบุคคลจะเข้าใจถึงทักษะและความรู้ในสมรรถนะผู้ประเมิน ดังในกรณีที่ผู้มีส่วนได้ส่วนเสียเน้นผู้เชี่ยวชาญทางเนื้อหา โดยละเลยความสำคัญของผู้เชี่ยวชาญทางการประเมิน อาจเกิดความลำเอียงในวิธีหรือรูปแบบที่แน่นอน ซึ่งอาจไม่สัมพันธ์กับคุณภาพหรือคุณค่าของสิ่งที่มุ่งประเมิน กลุ่มผู้มีส่วนได้ส่วนเสียหลายกลุ่ม (เช่น ผู้ให้ทุน พนักงานโครงการ ลูกค้า) อาจค้นหาคำประกอบของความเชี่ยวชาญที่น่าเชื่อถือที่แตกต่างกัน เหตุผลก็คือ ความน่าเชื่อถือของการประเมิน คือ มันใจได้ว่าความน่าเชื่อถือจะประนีประนอมไม่ให้เกิดปัญหาให้กับกลุ่มผู้มีส่วนได้ส่วนเสียและมุมมองของผู้มีส่วนได้ส่วนเสียที่ไม่คำนึงถึงความตรง

ประสิทธิผลต้นทุน (cost-effectiveness) เมื่อการประเมินคำนึงถึงความคุ้มค่าของการประเมินจึงมีความสำคัญขึ้นอยู่กับ ชนิดของค่าใช้จ่ายที่แตกต่างกันตามประเภทของผู้มีส่วนได้ส่วนเสีย ค่าใช้จ่ายที่เห็นชัดที่สุดเกี่ยวกับค่าจ้างที่จ่ายให้ทีมประเมิน แต่ค่าใช้จ่ายที่เห็นชัดน้อยที่สุด ได้แก่ ความคุ้มค่าของโอกาสในการเป็นส่วนหนึ่งของบุคลากรภายในการปฏิบัติงานประเมินในองค์กรตามหน้าที่ปกติ พื้นที่หรือสิ่งอำนวยความสะดวกและทรัพยากรอื่นๆ นอกจากนี้สิ่งที่สำคัญ คือ ค่าใช้จ่ายเกี่ยวกับเวลา หมายถึงการใช้เวลาในองค์กรที่สืบค้นสารสนเทศในการประเมิน เทียบกับกรอบเวลาที่กำหนดในการประเมิน การตอบคำถามในการประเมินเกี่ยวกับเหตุผลในการใช้จ่ายที่เป็นประโยชน์ต่อองค์กรและ/หรือ ชุมชน (เช่น ความรู้ที่เพิ่มขึ้น ความสามารถในการพัฒนาการเรียนรู้)

3. ประเภทของการประเมินอภิมาน

สุวิมล ว่องวาณิช (2550 อ้างถึงใน Bustelo, 2004) จัดประเภทของการประเมินอภิมานไว้ใน 4 มิติ คือ มิติแรก เกี่ยวกับบทบาทของการประเมินอภิมาน มิติที่สอง เป็นมิติเกี่ยวกับช่วงเวลา ของการประเมินอภิมาน มิติที่สามเป็นมิติเกี่ยวกับผู้ทำการประเมินอภิมาน และมิติที่สี่ เป็นมิติเกี่ยวกับเนื้อหาหรือระยะเวลาในการประเมินอภิมาน ดังต่อไปนี้

มิติแรก บทบาทของการประเมินอภิมาน (role of meta-evaluation) แบ่งออกเป็น 2 ประเภท คือ การประเมินอภิมานความก้าวหน้า (formative meta-evaluation) เป็นการประเมินในช่วงกำลังดำเนินการประเมินในส่วนของการวางแผนการประเมิน ช่วยให้นักประเมินวางแผน ทำการประเมิน ปรับปรุง ติความเกี่ยวกับสิ่งที่ประเมิน และการประเมินอภิมานผลสรุป (summative meta-evaluation) เป็นการประเมินหลังสิ้นสุดการประเมิน เพื่อช่วยให้ผู้ที่มีส่วนเกี่ยวข้องได้มองเห็นจุดแข็ง จุดอ่อนของการประเมิน และตัดสินคุณภาพ ตามมาตรฐานของการประเมิน

มิติที่สอง ช่วงเวลาของการดำเนินงาน (moment of meta-evaluation) แบ่งออกเป็น 2 ประเภทคือ การประเมินอภิมานก่อนมีกระบวนการประเมิน (ex-ante meta-evaluation) และการประเมินอภิมานระหว่างการดำเนินการประเมินหรือเมื่อการประเมินเสร็จสิ้น (ex-post facto meta-evaluation) ที่สอดคล้องกับ Schwandt และ Halpern (1998) ที่แบ่งเป็นการประเมินก่อน (provaluation) และการประเมินย้อนหลัง (retrovaluation)

มิติที่สาม หน่วยงานหรือบุคคลที่รับผิดชอบในการประเมินอภิมาน (agent in charge of meta-evaluation) แบ่งเป็น 2 ประเภท คือ การประเมินอภิมานภายใน (internal meta-evaluation) เป็นการประเมินอภิมานด้วยนักประเมินภายใน เพื่อควบคุมคุณภาพงานประเมินของตนเอง และการประเมินอภิมานภายนอก (external meta-evaluation) เป็นการประเมินอภิมานที่กระทำโดยทั้งนักประเมินภายในและนักประเมินภายนอก ซึ่งเป็นบุคคลที่ไม่ได้มีส่วนเกี่ยวข้องกับหน่วยงานนั้น เพื่อตรวจสอบกระบวนการประเมิน และสร้างความน่าเชื่อถือของผลการประเมิน

มิติที่สี่ เนื้อหาที่ทำการประเมินอภิมานหรือระยะเวลาของการประเมินอภิมาน (content or evaluation phase of meta-evaluation) แบ่งออกเป็น 3 ขั้นตอน คือ **ขั้นแรก** เป็นการประเมินอภิมานการออกแบบการประเมิน (design meta-evaluation) เป็นการประเมินอภิมานในส่วนของการกำหนดแนวคิด เหตุผล จุดมุ่งหมาย ผู้มีส่วนได้ส่วนเสีย การวางแผนการประเมิน เพื่อพัฒนาปรับปรุง และความรับผิดชอบต่อที่ตรวจสอบได้ **ขั้นที่สอง** เป็นการประเมินอภิมานในส่วนของการกระบวนการประเมิน (process meta-evaluation) เป็นขั้นตอนการนำแผนการประเมินไปปฏิบัติ เพื่อปรับแผนการประเมินให้มีประสิทธิผล **ขั้นสุดท้าย** เป็นการประเมินอภิมานในส่วนของการผลิต/ผลลัพธ์

จากการประเมิน (result meta-evaluation) เป็นการประเมินข้อสรุปของการประเมินเพื่อสะท้อนให้เห็นถึงความครบถ้วนสมบูรณ์ของข้อมูลในรายงานการประเมิน และความตรงต่อเวลาในการส่งผลการประเมิน

4. วิธีการและขั้นตอนที่ใช้ในการประเมินอภิमान

วิธีการและขั้นตอนที่ใช้ในการประเมินอภิमान ประกอบด้วย 2 หัวข้อ คือ 4.1) วิธีการที่ใช้ในการประเมินอภิमान 4.2) ขั้นตอนในการประเมินอภิमान ดังมีรายละเอียดต่อไปนี้

4.1 วิธีการที่ใช้ในการประเมินอภิमान สุวิมล ว่องวานิช (2550) ได้สรุปไว้ 2 แนวคิด คือ การประเมินอภิमानโดยใช้การตรวจเยี่ยมพื้นที่เชิงประเมิน (evaluative site visit meta-evaluation) และการประเมินอภิमानโดยการประเมินจากเอกสารรายงานการประเมิน (evaluation report based meta-evaluation) ซึ่งแนวคิดในการประเมินทั้ง 2 แนวคิดมีวิธีการในการประเมินแตกต่างกัน ดังต่อไปนี้ คือ การประเมินอภิमानโดยใช้การตรวจเยี่ยมพื้นที่เชิงประเมิน เป็นเทคนิคการประเมินที่มีการเก็บรวบรวมข้อมูลโดยใช้ประสบการณ์ของผู้ประเมินหรือจากรายงานของผู้อื่น การตรวจเยี่ยมต้องมีจุดเน้นเชิงประเมิน (evaluative focus) การตรวจเยี่ยมพื้นที่ประเมินมีข้อดี ก็คือ มักเก็บรวบรวมข้อมูลได้รวดเร็ว มีวิธีการเก็บรวบรวมข้อมูลได้หลากหลาย ทั้งการทบทวนเอกสาร การวิเคราะห์ สังเกต สัมภาษณ์ และจดบันทึก แต่ข้อจำกัดมักอาจสิ้นเปลืองงบประมาณ และหากผู้ประเมินไม่มีประสบการณ์ในการประเมิน อาจได้ข้อมูลไม่เพียงพอและไม่เข้าใจความซับซ้อนของพื้นที่ ส่วนแนวคิดการประเมินอภิमानโดยการประเมินจากเอกสารรายงานการประเมิน เป็นวิธีการตรวจสอบคุณภาพของการประเมินจากรายงานการประเมิน สามารถดำเนินการได้ง่าย ประหยัดเวลา แต่รายงานการประเมินจำเป็นต้องมีคุณภาพ มีข้อมูลหลักฐานอ้างอิง หรือแหล่งที่มาของข้อมูลที่ชัดเจนเพียงพอ ส่วนใหญ่ใช้ในการประเมินสรุปรวม ซึ่งหากใช้นักประเมินภายนอก ควรได้รับความร่วมมือจากนักประเมินปฐมภูมิภายในองค์กร ดังนั้นจึงสรุปได้ว่า หากงานประเมินขนาดใหญ่ ควรใช้วิธีการประเมินอภิमानภายนอกด้วยการตรวจเยี่ยมพื้นที่ จะเหมาะสมกว่า ซึ่งส่งผลกระทบต่อวงกว้าง และงบประมาณเพียงพอ ส่วนการประเมินอภิमानสำหรับงานประเมินขนาดเล็ก ใช้กระบวนการประเมินซ้ำแบบเดียวกัน ควรใช้การประเมินจากเอกสารรายงานจะเหมาะสมกว่า ดังเช่นการประเมินคุณภาพภายนอกสำหรับสถานศึกษา เพื่อเป็นการยกระดับมาตรฐานการประเมินภายนอก ควรใช้วิธีตรวจสอบคุณภาพจากการตรวจเยี่ยมพื้นที่เชิงประเมินเป็นระยะๆ แต่ทั้งนี้ต้องคำนึงระยะเวลาและงบประมาณในการประเมินด้วย

4.2 ขั้นตอนในการประเมินอภิमान

สำหรับขั้นตอนในการประเมินอภิमान สรุปได้ 9 ขั้นตอน ดังนี้ คือ 1) กำหนดจุดมุ่งหมายและประเภทของการประเมินอภิमान 2) กำหนดคณะนักประเมินอภิमान 3) กำหนดมาตรฐาน เกณฑ์ การตัดสินคุณภาพการประเมิน 4) การสร้างเครื่องมือ สำหรับการประเมินอภิमान 5) การรวบรวมข้อมูลสารสนเทศ 6) วิเคราะห์ข้อมูลสารสนเทศเชิงปริมาณและเชิงคุณภาพ 7) ตัดสินงานประเมินตามมาตรฐาน เกณฑ์การประเมิน 8) นำเสนอรายงานผลการประเมินอภิमान 9) สรุปผลการประเมินและประยุกต์ใช้ผลการประเมินอภิमान

Stufflebeam, Madaus, และ Kellagham (2000) ได้วิเคราะห์ขั้นตอนการประเมินอภิमानไว้ 10 ขั้นตอนดังต่อไปนี้

1) **ตัดสินและเตรียมการสร้างปฏิสัมพันธ์กับผู้มีส่วนได้ส่วนเสียในการประเมินอภิमान** โดยจัดการประชุมร่วมกันระหว่างนักประเมินอภิमान กับผู้มีส่วนได้ส่วนเสีย เป็นรายเดือน โดยอาจประชุมครั้งละประมาณ 2 ชั่วโมง และมีการเขียนสรุปรายงาน เพื่อให้ผู้ทรงคุณวุฒิอ่านเพื่อพิจารณาตัดสินวิเคราะห์จุดแข็งและจุดอ่อนในระบบการประเมิน

2) **กำหนดคุณสมบัติของทีมในการประเมินอภิमान** ให้มีความน่าเชื่อถือ โดยบุคคลในทีมประเมินจะต้องมีทั้งสมรรถนะในการประเมิน และได้รับความไว้วางใจจากผู้มีส่วนได้ส่วนเสีย ซึ่งบุคคลในทีมจะมีประสบการณ์ในการประเมินและเชี่ยวชาญในด้านที่แตกต่างกัน หัวหน้าทีมประเมินอภิमानจำเป็นต้องกำหนดประเภทการประเมินอภิमानและมีปฏิสัมพันธ์กับผู้มีส่วนได้ส่วนเสียอย่างเหมาะสม โดยมีคุณสมบัติที่เหมาะสมและได้รับการยอมรับจากทีมประเมิน

3) **นิยามคำถามในการประเมินอภิमान** ซึ่งคำถามในการประเมินต้องมีความชัดเจนสำหรับผู้รับผลการประเมินและผู้มีส่วนได้ส่วนเสีย เพื่อขยายข้อค้นพบให้ได้เข้าถึงความต้องการจำเป็นเกี่ยวกับสารสนเทศในการประเมินของผู้รับผลการประเมิน และประเมินเพื่อให้ได้ผลการประเมินตรงตามความต้องการ

4) **กำหนดมาตรฐานในการตัดสินระบบการประเมินหรือการประเมินเฉพาะเรื่องใดเรื่องหนึ่ง** เป็นการกำหนดมาตรฐานที่ใช้ในการประเมินอภิमान ได้แก่ มาตรฐานการประเมินโครงการ และมาตรฐานการประเมินบุคลากรของอเมริกา เป็นต้น และประยุกต์ใช้มาตรฐานการประเมินบุคลากร เช่น การประเมินระบบผู้บริหารของ GM (General Motors) เป็นต้น

5) **วางกรอบข้อตกลงในการประเมินอภิमान** เป็นการตกลงทำความเข้าใจและช่วยไม่ให้เกิดความเข้าใจผิดก่อน ระหว่างผู้รับผลการประเมินกับนักประเมินอภิमान และจัดเตรียมวิธีการแก้ปัญหาเบื้องต้นเกี่ยวกับการประเมิน การเขียนข้อตกลงสำหรับการประเมินอภิमानจะ

เปิดเผย แต่จะเป็นไปอย่างเหมาะสม โดยข้อตกลงขั้นสูงหมายถึง ความแตกต่างระหว่างความสำเร็จ และความล้มเหลวของการประเมินอภิमान

6) **รวบรวมและทบทวนความเหมาะสมของสารสนเทศที่เกี่ยวข้อง** หลังจากทำ ข้อตกลงระหว่างกันแล้ว นักประเมินอภิमानจำเป็นต้องทดสอบประเมินจากหลักฐานที่มีผลกระทบแน่ชัด เริ่มจากการรวบรวมและการประเมินสารสนเทศที่เปิดเผย การประเมินอภิमानบางครั้งเป็นเพียงสารสนเทศที่ใช้เพียงเพื่อสรุปผลการประเมินอภิमान เหตุผลที่ถูกต้องในการรวบรวมสารสนเทศเพิ่มเติม เนื่องจากสารสนเทศที่เปิดเผยยังไม่เพียงพอต่อการประเมิน หรือเพียงพอแต่ยังไม่น่าเชื่อถือในการตอบคำถามการประเมินอภิमान เมื่อสารสนเทศที่ได้รับการเปิดเผยได้รับการยอมรับพอ ในการเขียนรายงานการประเมินอภิमान การรวบรวมข้อมูลก็อาจเสียเวลาเปล่า

7) **รวบรวมสารสนเทศใหม่ๆ ที่จำเป็น เป็นข้อสรุป** ตัวอย่างเช่น บทสัมภาษณ์ ในพื้นที่จากผู้นำ และผู้ที่เกี่ยวข้อง ผลจากการสังเกตและการสำรวจ นอกจากนี้ก็อาจต้องทำการตรวจเยี่ยม เพื่อให้ได้ข้อมูลที่มีความตรงเชิงประเมินสำหรับเขียนรายงานการประเมินอภิमान

8) **วิเคราะห์ข้อค้นพบและตัดสินใจประเมินโดยยึดมาตรฐานการประเมิน** สารสนเทศที่ใช้ในการประเมินอภิमान ใช้การกระบวนการวิเคราะห์ทั้งเชิงปริมาณและเชิงคุณภาพ โดยใช้เทคนิคการนำเสนอด้วยกราฟเส้น กราฟแท่ง กราฟวงกลม และการวิเคราะห์เนื้อหาด้วยคอมพิวเตอร์ ความสำคัญในการวิเคราะห์ตัดสินในบริบทของมาตรฐานการประเมิน

9) **เตรียมและนำเสนอรายงานผลการประเมิน** หลังจากการวิเคราะห์สารสนเทศ ในการประเมินอภิमान นักประเมินอภิमानเตรียมและเสนอรายงานการประเมิน ในบางสถานการณ์ จะมีรายงานหลายฉบับ ตัวอย่างเช่น รายงานความสำเร็จของโครงการ รายงานประเมินระบบประเมิน บุคลากร รายงานการประเมินระบบการออกแบบ เป็นต้น โดยรายงานแต่ละฉบับต้องมีบทสรุปผู้บริหาร โดยทุกฉบับมีการนำเสนอรายงานทั้งฉบับร่าง และรายงานฉบับสมบูรณ์ต่อผู้รับผลการประเมิน

10) **ช่วยเหลือผู้รับผลการประเมินและผู้มีส่วนได้ส่วนเสียอื่น ในการตีความ และประยุกต์ใช้ผลการประเมินที่ค้นพบ** การรับรายงานการประเมินอภิमान ไม่ใช่เป็นการสิ้นสุดของการประเมินอภิमानที่สำเร็จ อีกบทบาทหนึ่งของนักประเมินอภิमानเป็นการส่งเสริมและช่วยเหลืออย่างเหมาะสมในการใช้ผลการประเมินอภิमानที่ค้นพบ เพื่อเป็นการวางรากฐานเริ่มต้นของการประเมินอภิमान

นอกจากนี้ สุวิมล ว่องวานิช (2551) ได้เสนอขั้นตอนของการประเมินอภิमानไว้ 6 ขั้นตอน ดังนี้

1) **การกำหนดจุดมุ่งหมายและประเภทของการประเมินอภิमान** เป็นขั้นตอนที่ผู้บริหาร หรือผู้มีอำนาจในการตัดสินใจ จะเป็นผู้กำหนดจุดมุ่งหมายของการประเมินอภิमानว่า เป็นการประเมินอภิमानเพื่อตรวจสอบคุณภาพการประเมินในระหว่างดำเนินการ เพื่อให้ผลการประเมินในการปรับปรุงการดำเนินงานในระยะต่อไป หรือเป็นการตรวจสอบคุณภาพการประเมินหลังสิ้นสุดการประเมิน เพื่อใช้ในการปรับปรุงวิธีการประเมินในครั้งต่อไป นอกจากนี้ต้องพิจารณาว่า การประเมินอภิमानจะประเมินโดยใช้บุคลากรในหน่วยงานตนเอง ระดับจุลภาค เป็นการประเมินอภิमानภายใน หรือบุคคลภายนอกที่ไม่มีส่วนเกี่ยวข้องหรือมีส่วนได้ส่วนเสียกับองค์กร ระดับมหภาค เป็นการประเมินอภิमानภายนอก ควรใช้นักประเมินอภิमानมืออาชีพ เพื่อจะช่วยลดอคติในการประเมิน สามารถความโปร่งใสในการดำเนินการที่สามารถตรวจสอบได้

2) **การกำหนดคณนักรประเมินอภิमान** นักประเมินอภิमान (meta evaluators) จะต้องเป็นบุคคลที่มีความรู้ ความสามารถ ประสบการณ์ในการประเมิน หรือเคยได้รับการอบรม ด้านการวิจัย และการประเมินมาก่อน โดยอาจเป็นนักประเมินมาก่อน หากเป็นนักประเมินอภิमान ความก้าวหน้า (formative meta evaluator) สามารถใช้นักประเมินในหน่วยงานได้เนื่องจากทราบบริบทการประเมินในองค์กรเป็นอย่างดี แต่หากเป็นนักประเมินสรุปรวม (summative meta evaluator) ควรใช้นักประเมินภายนอกองค์กรที่มีประสบการณ์ในการประเมินอภิमान มีความรู้ความสามารถในการประเมิน เพื่อชี้ให้เห็นจุดแข็ง จุดอ่อนของการประเมินได้เป็นอย่างดี

3) **การกำหนดเกณฑ์การตัดสินคุณภาพของการประเมิน** การประเมินอภิमानนิยมกำหนดเกณฑ์ในการตัดสินคุณภาพตามมาตรฐานการประเมินที่พัฒนาโดย Stufflebeam และคณะ ที่ประกอบด้วยมาตรฐานหลัก 4 ด้าน คือ ด้านอรรถประโยชน์ ด้านความเป็นไปได้ ด้านความเหมาะสมชอบธรรม และด้านความถูกต้อง

4) **การสร้างเครื่องมือสำหรับการประเมินอภิमान** เครื่องมือที่ใช้ในการประเมินอภิमानที่พัฒนาขึ้น ต้องสอดคล้องกับเกณฑ์ มาตรฐานการประเมินที่ใช้ นอกจากนี้ต้องมีความตรงกับการประเมิน รูปแบบของเครื่องมือต้องใช้งานได้ง่าย สะดวก และตรวจสอบคุณภาพของการประเมินได้

5) **การตรวจสอบคุณภาพของการประเมิน** เป็นขั้นตอนของการนำเครื่องมือสำหรับประเมินอภิमानไปเก็บรวบรวมข้อมูล ตามร่องรอยหลักฐานการดำเนินการประเมิน ตัดสินคุณภาพการประเมินตาม เกณฑ์ มาตรฐานการประเมินอภิमानที่กำหนด ให้คะแนนตามระดับ

คุณภาพแต่ละข้อ ทำการวิเคราะห์ประมวลผลข้อมูล และสรุปผลค่าคะแนนรายข้อ และภาพรวมของการประเมิน

6) การสรุปผลการประเมินอภิमानและการนำเสนอผลต่อผู้ที่เกี่ยวข้อง การสรุปผลการประเมินอภิमानจะต้องชี้ให้เห็นจุดเด่น จุดที่ต้องปรับปรุงแก้ไข ในด้านวิธีการประเมิน และคุณภาพของรายงานผลการประเมิน เพื่อให้ผลการประเมินมีความน่าเชื่อถือมากยิ่งขึ้น หากเป็นการประเมินอภิमानสรุปรวมให้นำเสนอจุดที่บกพร่องเพื่อใช้ในการปรับปรุงการประเมินในครั้งต่อไป ควรระบุประเด็นที่ผู้บริหาร หรือผู้มีอำนาจในการตัดสินใจนำผลการประเมินไปใช้ในการวางแผนและพัฒนาการดำเนินงานอย่างน่าเชื่อถือต่อไปได้

สรุปแนวคิดของนักวิชาการ ได้แก่ Stufflebeam, Madaus, และ Kellagham (2000) และสุวิมล ว่องวาณิช (2551) ได้ดังตารางที่ 2.4

ตารางที่ 2.4 สรุปขั้นตอนการประเมินอภิमान

Stufflebeam, Madaus, และ Kellagham (2000)	สุวิมล ว่องวาณิช (2551)	ความหมาย
1) ตัดสินและเตรียมการสร้างปฏิสัมพันธ์กับผู้มีส่วนได้ส่วนเสียในการประเมินอภิमान	-	จัดการประชุมร่วมกันระหว่างนักประเมินอภิमान กับผู้มีส่วนได้ส่วนเสีย เพื่อพิจารณาตัดสินวิเคราะห์จุดแข็งและจุดอ่อนในระบบการประเมิน
2) กำหนดคุณสมบัติของทีมในการประเมินอภิमान	การกำหนดคณะนักประเมินอภิमान	ทีมประเมินจะต้องมีความรู้ความสามารถ ประสบการณ์ในการประเมิน และได้รับความไว้วางใจจากผู้มีส่วนได้ส่วนเสีย
3) นิยามคำถามในการประเมินอภิमान	การกำหนดจุดมุ่งหมายและประเภทของการประเมินอภิमान	การกำหนดคำถามและจุดมุ่งหมายในการประเมิน เพื่อให้ได้ผลการประเมินตรงตามความต้องการ และตรวจสอบคุณภาพการประเมินในระหว่างดำเนินการ และสิ้นสุดการประเมิน เพื่อปรับปรุงการดำเนินงานในระยะต่อไป
4) กำหนดมาตรฐานในการตัดสินระบบการประเมิน	การกำหนดเกณฑ์การตัดสินคุณภาพของการประเมิน	เป็นการกำหนดมาตรฐานที่ใช้ในการประเมินอภิमान ได้แก่ Stufflebeam ที่ประกอบด้วยมาตรฐานหลัก 4 ด้าน

ตารางที่ 2.4 (ต่อ)

Stufflebeam, Madaus, และ Kellaghan (2000)	สุวิมล ว่องวานิช (2551)	ความหมาย
5) วางกรอบข้อตกลงในการประเมิน อภิमान	-	เป็นเขียนข้อตกลง ระหว่างผู้รับผลการ ประเมินกับนักประเมินอภิमान และ จัดเตรียมวิธีการแก้ปัญหาเบื้องต้น เกี่ยวกับการประเมิน
6) รวบรวมและทบทวนความเหมาะสม ของสารสนเทศที่เกี่ยวข้อง	การสร้างเครื่องมือสำหรับการ ประเมินอภิमान	การรวบรวมสารสนเทศจาก การ ประเมิน เพื่อนำไปสู่การสรุปผล ประเมินอภิमान โดยสร้างเครื่องมือที่มี คุณภาพ เพื่อรวบรวมสารสนเทศที่ นำเชื่อถือในการตอบคำถามการ ประเมินอภิमान
7) รวบรวมสารสนเทศใหม่ๆ ที่จำเป็น เป็นข้อสรุป	การตรวจสอบคุณภาพของ การประเมิน	นำเครื่องมือไปเก็บรวบรวมสารสนเทศ จากการสัมภาษณ์ การสังเกตและการ สำรวจ เพื่อให้ได้ข้อมูลที่มีความตรง เชิงประเมิน
8) วิเคราะห์ข้อค้นพบและตัดสินใจ ประเมิน	การสรุปผลการประเมิน อภิमानและการนำเสนอผล ต่อผู้ที่เกี่ยวข้อง	ใช้การกระบวนการวิเคราะห์ที่ทั้งเชิง ปริมาณและเชิงคุณภาพ สรุปผลการ ประเมินอภิमान เพื่อชี้ให้เห็นจุดเด่น จุดที่ต้องปรับปรุงแก้ไข เพื่อให้ผลการ ประเมินมีความน่าเชื่อถือมากยิ่งขึ้น
9) เตรียมและนำเสนอรายงานผลการ ประเมิน		
10. การตีความและประยุกต์ใช้ผลการ ประเมินที่ค้นพบ	-	เป็นการส่งเสริมและช่วยเหลืออย่าง เหมาะสมในการใช้ผลการประเมิน อภิमानที่ค้นพบ เพื่อเป็นการ วางรากฐานของการประเมินอภิमान

Davidson (2005) ได้เสนอวิธีการประเมินอภิमानเพื่อเป็นทางเลือกในการทำการ
ประเมินอภิमानไว้ 3 ทางเลือก ดังนี้ **ทางเลือกแรก** เป็นการประเมินตามมาตรฐานการประเมิน
โครงการ ของ Joint Committee on Standards for Educational Evaluation (1994) โดยทำแบบ
ตรวจสอบรายการประเมินอภิमान (meta evaluation checklist) มีประโยชน์ในการใช้ตรวจสอบ
รายการตามจุดเน้น 4 ด้าน คือ ด้านความเหมาะสม (propriety) ความเป็นไปได้ (feasibility)
อรรถประโยชน์ (utility) และด้านความถูกต้อง (accuracy) **ทางเลือกที่สอง** เป็นวิธีการประเมินโดย
ใช้ทีมประเมินที่แตกต่างกัน ซึ่งมีความเป็นอิสระของการประเมิน เป็นการทำการประเมินในสิ่ง
เดียวกัน และมีการรายงานเพื่อเปรียบเทียบผลการประเมิน พิจารณาลักษณะที่เหมือนกันและหากพบสิ่ง

แตกต่างกัน จะค้นหาข้อสรุปเดียวกันหรือข้อค้นพบที่ถูกต้อง **ทางเลือกสุดท้าย** เป็นวิธีการผสม โดยวิธีนี้ ผู้ประเมินอภิमानจะทำการประเมิน และมีการตรวจสอบไขว้ผลจากข้อค้นพบที่น่าสงสัย ซึ่งในการประเมินแบบดั้งเดิมผู้มีส่วนได้ส่วนเสียอาจมีข้อสงสัยในผลการประเมินที่ค้นพบ จึงต้องมีการประเมินหรือตรวจสอบข้อสงสัยดังกล่าว

สำหรับเทคนิควิธีการที่ใช้ในการประเมินอภิमान สรุปจากการสังเคราะห์งานวิจัยที่เกี่ยวข้องได้ 3 วิธีหลัก ซึ่งสามารถสรุปจัดกลุ่มเทคนิควิธีการประเมินอภิमानได้ดังตารางที่ 2.5 โดยมีรายละเอียดต่อไปนี้ คือ

- 1) การสัมภาษณ์ผู้ที่เกี่ยวข้องทั้งสัมภาษณ์โดยตรง และการสัมภาษณ์ทางโทรศัพท์ (Madaus, & Mcsherry, 2007, Oxfam Australia, 2006)
- 2) การทำสนทนากลุ่ม (focus group) จากสมาชิกในกลุ่ม (Madaus, & Mcsherry, 2007)
- 3) การสังเคราะห์เอกสาร/รายงานวิจัย/ รายงานการประเมิน รวมถึงการทบทวนเอกสารฐานข้อมูล (Scott-little, Hamann, & Jurs, 2002; Oxfam Australia, 2006; Windsor, Boyd, & Orleans, 1998; Bornmann, Mittag & Daniel, 2006; Hartmann & Loizides, 2001)

ตารางที่ 2.5 สรุปเทคนิควิธีการประเมินอภิमानจากงานวิจัยที่เกี่ยวข้อง

ชื่อนักวิจัย	นักประเมิน	วิธีการประเมิน	การเก็บรวมข้อมูล
Madaus, และ Mcsherry (2007)	พนักงานของสำนักงานส่งเสริม สนับสนุน ดูแล การมองข้ามเด็ก (OCAIO)	เทคนิคบูรณาการ และความแม่นยำของวิธีการประเมิน	-การสัมภาษณ์ พระ พระสงฆ์และผู้นำศาสนา ผู้บริหารโรงเรียน และผู้นำการศึกษา -การสัมภาษณ์ทางโทรศัพท์ จากผู้ปกครองเด็กในโรงเรียน และผู้ปกครองเด็กในศาสนา -การจัดทำสนทนากลุ่ม จากครูในโรงเรียน ครูสอนศาสนา สมาชิกของทีม และอาสาสมัคร
Oxfam Australia (2006)	โดยนักประเมินภายในและภายนอก	การศึกษาประเมินในฐานข้อมูล Oxfam	-การทบทวนเอกสาร ทบทวนฐานข้อมูลการประเมิน -การสัมภาษณ์พนักงาน -การประเมินเชิงลึก -การสังเคราะห์ข้อค้นพบจากรายงานการประเมินย่อย
Scott-little, Hamann, และ Jurs (2002)	นักวิจัย	วิธีการและการสังเคราะห์ข้อค้นพบแบบบรรยาย	การสังเคราะห์จากงานประเมินกิจกรรมหลังการเรียนในชั้นเรียน

ตารางที่ 2.5 (ต่อ)

ชื่อนักวิจัย	นักประเมิน	วิธีการประเมิน	การเก็บรวมข้อมูล
Windsor, Boyd, และ Orleans (1998)	นักวิจัย	-ประเมินคุณภาพงานประเมิน ประเมิน -สังเคราะห์ผลการประเมินโครงการ	-การประเมินคุณภาพงานประเมินสอดคล้องกับ มาตรฐานการประเมินของ Joint Committee -สังเคราะห์ข้อมูลจากงานวิจัยโครงการ เล็กสู่นานหรือในสตรีตั้งครรภ์
Bornmann, Mittag & Daniel (2006)	ผู้ที่เกี่ยวข้อง	กระบวนการประเมินแบบหลายขั้นตอน	ประเมินคุณภาพการประเมินการสอนและการเรียนในสถาบันอุดมศึกษา
Hartmann and Loizides (2001)		-มาตรฐาน Joint Committee -มาตรฐานสำหรับสำรวจอินเตอร์เน็ต	-แบบสำรวจความก้าวหน้าทางเทคโนโลยี (ATE) ทดสอบจากส่วนหนึ่งของแบบสำรวจจากรายงานของศูนย์ เกี่ยวกับผลกระทบและประสิทธิผลของโครงการ ATE

สรุปได้ว่า การประเมินอภิมาน เป็นการประเมินงานประเมิน เพื่อควบคุม ตรวจสอบ และประเมินคุณภาพของงานประเมิน โดยการประเมินอภิมาน ประกอบด้วย 3 ขั้นตอนหลัก คือ 1) กำหนดวัตถุประสงค์ในการประเมินอภิมานงานประเมิน 2) ประเมินงานประเมินเพื่อพิจารณาตัดสินคุณค่าของงานประเมินตามเกณฑ์มาตรฐานที่น่าเชื่อถือ 3) นำผลการประเมินอภิมานไปใช้ประโยชน์ ทั้งนี้แนวคิดของ Stufflebeam และ Davidson ที่สอดคล้องกันในเรื่องของความน่าเชื่อถือของผู้ประเมินอภิมาน และเกณฑ์มาตรฐานที่ใช้ในการประเมินอภิมาน

5. การสังเคราะห์มาตรฐานการประเมินอภิมาน

การประเมินอภิมานแบ่งออกเป็น 2 วิธี คือ การประเมินอภิมานโดยการตรวจเยี่ยมเชิงพื้นที่ และการประเมินอภิมานจากเอกสารรายงานการประเมิน ซึ่งการประเมินอภิมานที่นิยมใช้ในปัจจุบัน เป็นการประเมินจากรายงานการประเมิน มากกว่าการประเมินกระบวนการประเมิน หรือการประเมินระหว่างดำเนินการประเมิน (formative meta-evaluation) ซึ่งส่วนใหญ่แล้วการประเมินอภิมานจะนิยมใช้ แบบตรวจสอบสำหรับการประเมินอภิมานรายงานประเมินโครงการ (program evaluations meta-evaluation checklist) ที่พัฒนาขึ้นโดย Stufflebeam (1999) ใน 4 ด้าน คือ ธรรมชาติประโยชน์ (utility) ความถูกต้อง (accuracy) ความเป็นไปได้ (feasibility) และความเหมาะสม (propriety) จาก Program Evaluation Standards ของ the Joint Committee on Standard for Educational Evaluation (1994) เป็นหลัก เพื่อประเมินจุดแข็ง จุดอ่อนของการประเมิน แต่การประเมินอภิมานจากรายงานการประเมินเป็นการสรุปข้อมูลที่มีในรายงานหรือได้จากรวบรวมข้อมูลจากผู้มีส่วนได้ส่วนเสีย

เป็นหลัก ดังนั้นจึงมีความจำเป็นต้องกำหนดมาตรฐาน เกณฑ์ในการประเมินที่มีความยุติธรรม ซึ่ง Davidson (2005) ได้กล่าวว่า กลยุทธ์ในการประเมินงานประเมินจากเอกสารรายงาน ควรพิจารณาจาก 5 มาตรฐานหลักหมายถึง การประเมินต้องทำอย่างถูกต้องและสรุปอย่างสมเหตุสมผลมีประโยชน์ ต่อผู้รับการประเมินและผู้ที่เกี่ยวข้อง ทำการประเมินอย่างมีจริยธรรม ตามกฎระเบียบ เชี่ยวชาญและเหมาะสม น่าเชื่อถือต่อผู้ที่เกี่ยวข้อง รวมถึงประหยัด รวดเร็ว และไม่ยุ่งยาก ดังได้แก่ 1) ความตรงของการประเมิน (validity) 2) รรถประโยชน์ (utility) 3) จริยธรรมในการประเมิน (ethicality) 4) ความน่าเชื่อถือของผู้ประเมิน (credibility) และ 5) ประสิทธิภาพต้นทุน (cost-effectiveness)

สำหรับการสังเคราะห์มาตรฐานการประเมินอภิมานในครั้งนี้ ทำการสังเคราะห์โดยการศึกษาดูเอกสาร งานวิจัยที่เกี่ยวข้องกับมาตรฐานการประเมิน และการประยุกต์ใช้มาตรฐานการประเมิน ใน 7 แนวทางหลักๆ คือ มาตรฐานการประเมิน Program Evaluation Standards (PES) ตามแนวทางของ the Joint Committee on Standard for Educational Evaluation (1994) และการประยุกต์ใช้มาตรฐานการประเมินในบริบทของประเทศต่างๆ ได้แก่ 1) The African Evaluation guidelines (2002) (AEG) ประเทศแอฟริกา 2) SEVAL's Evaluation Standards (2004) (SEVAL) ประเทศสวีเดน และ 3) The DeGEval's Standards (2003) (DeGEval) ประเทศเยอรมัน นอกจากนี้พิจารณามาตรฐานของผู้ประเมิน ได้แก่ 1) Guiding Principles for Evaluators ของ American Evaluation Association [AEA] (2003) 2) Essential Skills Series in Evaluation ของ Canadian Evaluation Society (1999) (CES) และ 3) งานวิจัย Essential Competencies for Program Evaluators (Ghere & et al., 2006) (ECPE) เพื่อเป็นแนวทางในการพัฒนาตัวบ่งชี้ในการประเมินอภิมาน ดังรายละเอียดต่อไปนี้

5.1 มาตรฐานความตรงของการประเมิน (validity) เป็นการตัดสินลงข้อสรุปผลการประเมินอย่างถูกต้องสมเหตุสมผล โดยพิจารณาข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพ เปรียบเทียบกับเกณฑ์ ประกอบด้วย 11 ตัวบ่งชี้ ดังตารางที่ 2.6

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2.6 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานความตรงของการประเมิน

Validity	PES	AEG	SEVAL	DeGEval	AEA	CES
	1994	2002	2004	2003	2003	1999
1. การจัดการระบบเอกสารและกำหนดรูปแบบคำถามในการประเมิน	✓	✓	✓	✓		✓
2. การวิเคราะห์บริบทของงานประเมิน	✓	✓	✓	✓		
3. การนิยามวัตถุประสงค์ เงื่อนไข และออกแบบการประเมิน	✓	✓	✓	✓		✓
4. การใช้วิธีการประเมินที่หลากหลาย	-				✓	
5. การใช้แหล่งสารสนเทศที่น่าเชื่อถือ	✓	✓	✓	✓	✓	
6. สารสนเทศที่ได้มีความเที่ยงและ/หรือความตรง	✓	✓	✓	✓		✓
7. การบริหารข้อมูลและระบบของสารสนเทศ	✓	✓	✓	✓	✓	✓
8. การวิเคราะห์สารสนเทศเชิงปริมาณและ/หรือเชิงคุณภาพ	✓	✓	✓	✓	✓	✓
9. การลงข้อสรุปที่มีเหตุผลเหมาะสมถูกต้อง	✓	✓	✓	✓		
10. การรายงานผลอย่างเป็นไม่มีอคติ	✓	✓	✓	✓		
11. การประเมินอภิमान	✓	✓	✓	✓		

5.2 มาตรฐานอรรถประโยชน์ (utility) เป็นการทำให้เกิดความมั่นใจได้ว่าการประเมินจะใช้สารสนเทศที่จำเป็น ตรงกับเป้าหมายของผู้ใช้ผลการประเมิน ซึ่งการประเมินที่ดีจะเป็นประโยชน์ต่อผู้รับผลการประเมิน ประกอบด้วย 9 ตัวบ่งชี้ ดังตารางที่ 2.7

ตารางที่ 2.7 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานด้านอรรถประโยชน์

Utility	PES	AEG	SEVAL	DeGEval	AEA	CES
	1994	2002	2004	2003	2003	1999
1. การระบุผู้มีส่วนได้ส่วนเสีย	✓	✓	✓	✓		✓
2. ขอบเขตและการเลือกการเก็บรวบรวมข้อมูล	✓	✓	✓	✓	✓	✓
3. การระบุคุณค่าของการประเมิน	✓	✓	✓	✓		✓
4. การประเมินตามข้อจำกัดของทรัพยากร งบประมาณ อุปสรรคในการประเมิน						✓
5. การพัฒนาการวัดผลลัพธ์และเกณฑ์ที่บ่งชี้ถึงความสำเร็จ						✓
6. ความชัดเจนของรายงานการประเมิน	✓	✓	✓	✓	✓	
7. การวิเคราะห์และการรายงานผลการประเมินภายในเวลาที่กำหนด พร้อมทั้งเผยแพร่	✓	✓	✓	✓	✓	✓
8. การใช้ประโยชน์จากการประเมิน						✓
9. ผลกระทบของงานประเมิน	✓	✓	✓	✓		✓

5.3 มาตรฐานจริยธรรมในการประเมิน (ethicality) เป็นการประเมินที่สมบูรณ์และยุติธรรมต่อผู้รับผลการประเมิน คำนึงถึงสิทธิมนุษยชน และประโยชน์สาธารณะ ประกอบด้วย 9 ตัวบ่งชี้ ดังตารางที่ 2.8

ตารางที่ 2.8 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานจริยธรรมในการประเมิน

Ethicality	PES	AEG	SEVAL	DeGEval	AEA	CES
	1994	2002	2004	2003	2003	1999
1. กระบวนการปฏิบัติ เน้นการสร้างกรอบแนวทาง และวิธีการประเมิน	✓	✓	✓	✓		✓
2. การเป็นที่ยอมรับทางการเมือง	✓	✓	✓	✓		
3. การมุ่งเน้นความสำคัญของการให้บริการ ปรับปรุงคุณภาพอย่างต่อเนื่อง	✓	✓				✓
4. การทำข้อตกลงอย่างเป็นทางการในการออกแบบ การประเมิน	✓	✓	✓	✓		✓
5. การปกป้องสิทธิมนุษยชนของตัวอย่าง โดยยึด มาตรฐานและคุณธรรมของอาชีพ	✓	✓	✓	✓	✓	
6. การมีปฏิสัมพันธ์กับผู้ที่เกี่ยวข้อง	✓	✓	✓		✓	✓
7. การประเมินที่ยุติธรรมและสมบูรณ์	✓	✓	✓	✓		✓
8. การเปิดเผยผลการประเมิน	✓	✓	✓	✓		
9. คำเนิ่งถึงผลประโยชน์ที่เกี่ยวข้องกับสาธารณะ และสวัสดิภาพของสังคม					✓	

5.4 มาตรฐานความน่าเชื่อถือของผู้ประเมิน (credibility) เป็นการทำการประเมิน เพื่อตรวจสอบสมรรถนะที่จำเป็นของผู้ประเมิน เพื่อความน่าเชื่อถือของข้อค้นพบที่ได้รับจากการ ประเมิน ประกอบด้วย 2 ตัวบ่งชี้ ดังตารางที่ 2.9

ตารางที่ 2.9 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานด้านความน่าเชื่อถือของการประเมิน

Credibility	PES	AEG	SEVAL	DeGEval	AEA	ECPE
	1994	2002	2004	2003	2003	2006
1. สมรรถนะของผู้ประเมิน	✓	✓	✓	✓	✓	
1.1 สมรรถนะด้านความรู้ เทคนิคและความ ชำนาญในการประเมิน					✓	✓
1.2 สมรรถนะด้านมาตรฐานคุณธรรม จริยธรรม ของผู้ประเมิน					✓	✓
1.3 สมรรถนะด้านการวิเคราะห์ การตัดสินใจ						✓
1.4 สมรรถนะด้านทักษะในการบริหารจัดการ โครงการ					✓	✓
1.5 สมรรถนะด้านทักษะในการสื่อสาร การแก้ปัญหาของผู้ประเมิน					✓	✓
2. ความเป็นอิสระยุติธรรม ไม่มีผลประโยชน์ทับซ้อน	✓	✓	✓			

5.5 มาตรฐานประสิทธิผลต้นทุน (cost-effectiveness) เป็นการประเมินที่คำนึงถึงความคุ้มค่าของค่าใช้จ่าย และต้นทุนที่ใช้ในการประเมิน ประกอบด้วย 2 ตัวบ่งชี้ ดังตารางที่ 2.10

ตารางที่ 2.10 ผลการสังเคราะห์ตัวบ่งชี้ ในมาตรฐานประสิทธิผลต้นทุน

Cost-effectiveness	PES 1994	AEG 2002	SEVAL 2004	DeGEval 2003	AEA 2003	CES 1999
1. ประสิทธิภาพของต้นทุนการประเมิน	✓	✓	✓	✓	✓	✓
1.1 ด้านการเงิน						
1.2 ด้านบุคลากร						
1.3 ด้านสิ่งอำนวยความสะดวก						
1.4 ด้านทรัพยากรอื่นๆ						
2. ความรับผิดชอบในเรื่องการเงิน	✓	✓				

การสังเคราะห์ตัวบ่งชี้ในมาตรฐานการประเมินอภิมานตามมาตรฐานของ Davidson (2005) ที่ครอบคลุมทุกมาตรฐานการประเมิน program evaluation standard และมาตรฐานอื่นๆที่เกี่ยวข้อง ใน 7 แนวทางการประเมิน สรุปได้ดังต่อไปนี้ คือ 1) มาตรฐานความตรงของการประเมิน ประกอบด้วย 11 ตัวบ่งชี้ ที่สังเคราะห์มาจากมาตรฐานการประเมิน 6 มาตรฐาน ได้แก่ PES, AEG, SEVAL, DeGEval, AEA, และ CES 2) มาตรฐานอรรถประโยชน์ ประกอบด้วย 9 ตัวบ่งชี้ ที่สังเคราะห์มาจากมาตรฐานการประเมิน 6 มาตรฐาน ได้แก่ PES, AEG, SEVAL, DeGEval, AEA, และ CES 3) มาตรฐานจริยธรรมในการประเมิน ประกอบด้วย 9 ตัวบ่งชี้ ที่สังเคราะห์มาจากมาตรฐานการประเมิน 6 มาตรฐาน ได้แก่ PES, AEG, SEVAL, DeGEval, AEA, และ CES เช่นเดียวกัน 4) มาตรฐานความน่าเชื่อถือของนักประเมิน ประกอบด้วย 2 ตัวบ่งชี้ ที่สังเคราะห์มาจากมาตรฐานการประเมิน 6 มาตรฐาน ได้แก่ PES, AEG, SEVAL, DeGEval, AEA, และ ECPE และ 5) มาตรฐานประสิทธิผลต้นทุน ประกอบด้วย 2 ตัวบ่งชี้ ที่สังเคราะห์มาจากมาตรฐานการประเมิน 6 มาตรฐาน ได้แก่ PES, AEG, SEVAL, DeGEval, AEA, และ CES

6. งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องแบ่งออกเป็น 3 กลุ่ม คือ 6.1) การประเมินอภิมานตามมาตรฐานการประเมิน The Program Evaluation Standards และมาตรฐานที่พัฒนาขึ้นตามแนวทางของ The Joint Committee on Standards for Educational Evaluation ได้แก่ งานวิจัยของ Hartmann และ Loizides (2001) และ Patel (2002) 6.2) การประเมินอภิมานงานประเมิน หรือเทคนิควิธีการที่ใช้ในการประเมินอภิมาน ได้แก่ งานวิจัยของ Madaus และ Mcsherry (2007) Oxfam Australia (2006), Scott-little, Hamann, และ Jurs (2002) Widsor, Boyd, และ Orleans

(1998) Bjornklide และ Bason (2000) สมาคมครุศาสตร์อุตสาหกรรมแห่งประเทศไทย (2547) วารุณี ลัภนโชคดี (2550) และศูนย์ส่งเสริมและพัฒนาพลังแผ่นดินเชิงคุณธรรม (ศูนย์คุณธรรม) สำนักงานบริหารและพัฒนาองค์ความรู้ (องค์การมหาชน) (2551) 6.3) การประเมินอภิมานจาก ผลงานการเรียน ได้แก่ งานวิจัยของอมรรัตน์ พันธุ์งาม (2543) และภัทราวรรณ ภิรมย์วงศ์ (2546)

6.1 การประเมินอภิมานตามมาตรฐานการประเมิน The Program Evaluation Standards และมาตรฐานที่พัฒนาขึ้นตามแนวทางของ The Joint Committee on Standards for Educational Evaluation ได้แก่ Patel (2002) ซึ่งได้ทำการพัฒนาเป็นมาตรฐานการประเมิน The African Evaluation Guidelines (AEGs) เป็นแนวทางในการประเมิน ส่วนงานวิจัยของ Hartmann and Loizides (2001) ได้ใช้มาตรฐาน The Program Evaluation Standards (2nd ed.) ของ The Joint Committee on Standards for Educational Evaluation ในการประเมินอภิมาน Web-Based ด้วยการประเมินเชิงสำรวจของ ATE ซึ่งทำการสำรวจโดย ศูนย์การประเมินในมูลนิธิวิทยาศาสตร์แห่งชาติ (NSF) การศึกษาความก้าวหน้าทางเทคโนโลยี (ATE) ประยุกต์ใช้ 2 มาตรฐาน คือ 1) The Program Evaluation Standards (2nd ed.) ของ The Joint Committee on Standards for Educational Evaluation ที่ทำการประเมินในด้านอรรถประโยชน์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และความถูกต้อง (accuracy) และ 2) เป็นมาตรฐานสำหรับสำรวจ อินเทอร์เน็ต ได้มาจากแบบสำรวจเมลล์และอินเทอร์เน็ตของ Don Dillman ด้วยแบบสำรวจ ATE ทดสอบจากส่วนหนึ่งของแบบสำรวจจากรายงานของศูนย์ เกี่ยวกับผลกระทบและประสิทธิผลของ โครงการการศึกษาความก้าวหน้าทางเทคโนโลยี (ATE) ข้อมูลจากการศึกษาผลการไม่ตอบ แบบสอบถามในปี ค.ศ. 2000 รายงานการตรวจเยี่ยม สารสนเทศของ ATE web site กำหนดการ ดำเนินการ กระบวนการรับรองข้อมูล เอกสารที่เป็นประโยชน์ที่หาได้จากการตอบแบบสำรวจ emails ที่ส่งในช่วงระหว่างการสำรวจปี ค.ศ. 2001 ในรูปอิเล็กทรอนิกส์ และemail ให้ข้อมูลป้อนกลับ นอกจากนี้ยังสัมภาษณ์พนักงานของศูนย์การประเมิน 2 ครั้ง เพื่อเป็นสารสนเทศเพิ่มเติม การสำรวจ web-based ประเมินจากจำนวนเวลาที่เข้าใช้ การใช้ web-browsers การเข้าถึง the ATE Survey site ตามขั้นตอนอัตโนมัติ ผลการศึกษาพบว่า ทั้งมาตรฐานการประเมินโครงการ และแบบสำรวจตาม แนวทางของ Dillman นั้น แบบสำรวจ ATE เป็นเครื่องมือการประเมินที่ดี ซึ่งใช้ได้ถูกเวลา ทั่วถึง การ รายงานเป็นประโยชน์ต่อ NSF และต่อโครงการ ATE แต่ควรระมัดระวังและทบทวนการปฏิบัติการ สำรวจโดยพนักงานของศูนย์การประเมินในเรื่องของหลักฐาน และผลการประเมินเพื่อการปรับปรุงที่ยั่งยืนควรทำการประเมินซ้ำครั้งที่สอง แม้การประเมินในครั้งมีผลทางบวกสูงก็ตาม จึงเสนอแนะเพื่อ เป็นประโยชน์ในการปรับปรุงกลไก และการใช้แบบสำรวจอย่างเป็นมิตร

สำหรับงานวิจัยของ Patel (2002) ทำการศึกษาประเมินอภิमानอิงหลักการประเมินของแอฟริกัน เพื่อควบคุมคุณภาพของกระบวนการประเมิน โดยพิจารณาจากแนวปฏิบัติที่ดี (good practice) โดยสมาคมประเมินของแอฟริกันและเครือข่าย The African Evaluation Guidelines (AEGs) มาตรฐาน 30 ข้อ ที่ปรับปรุงมาจาก Program Evaluation Standard ของ Joint Committee on Standards for Educational Evaluation ใช้ checklist ประเมินควบคุมคุณภาพของงานประเมิน 14 ชั้นที่เผยแพร่ใน Evaluation and Program Planning โดย AEGs ประกอบด้วย 4 ด้าน คือ วัตถุประสงค์ (utility) ความเป็นไปได้ (feasibility) ความเหมาะสม (propriety) และ ความถูกต้อง (accuracy) ด้วยมาตรวัด 5 ระดับ ใช้ผู้ประเมิน 3 คน พบว่ามีคะแนนอยู่ในช่วง 1-4 คะแนน โดยข้อที่มีคะแนนมัธยฐาน 4 คะแนน ได้แก่ การระบุผู้มีส่วนได้ส่วนเสีย (stakeholder identification) ความชัดเจนของรายงานการประเมิน (report clarity) ผลกระทบของงานประเมิน (evaluation impact) กระบวนการปฏิบัติ (Practical Procedures) ประสิทธิภาพของต้นทุนการประเมิน (cost effectiveness) การมุ่งเน้นความสำคัญของการให้บริการ (service orientation) การปกป้องสิทธิมนุษยชนของผู้มีส่วนได้ส่วนเสีย (rights of stakeholder) การจัดการระบบเอกสารของโครงการ (program documentation) การวิเคราะห์บริบทของโครงการ (context analysis) การบรรยายวัตถุประสงค์และกระบวนการประเมิน (described purposes and procedures) และการวิเคราะห์สารสนเทศเชิงปริมาณอย่างมีประสิทธิภาพ (analysis of quantitative information) สำหรับข้อที่เป็นจุดอ่อน มีคะแนนมัธยฐาน 2 คะแนน คือ การระบุคุณค่า (values identification) โดยการประเมินอภิमानสามารถช่วยให้สามารถบ่งชี้จุดอ่อน จุดแข็งของการวิจัยเพื่อเผยแพร่ในสาขาเดียวกัน ในกลุ่มสาขาวิชานั้นได้ และนอกจากนี้ยังสามารถช่วยคัดเลือกเอกสารสำหรับเผยแพร่ได้อีกด้วย

6.2 การประเมินอภิमानงานประเมิน งานวิจัยส่วนมากเป็นการประเมินคุณภาพงานประเมินและสังเคราะห์งานประเมิน (Widsor, Boyd, and Orleans, 1998; Bjornklide and Bason, 2000; Oxfam Australia, 2006; Madaus และ Mcsherry, 2007; สมาคมครุศาสตร์อุตสาหกรรมแห่งประเทศไทย, 2547; วารุณี ถิ่นโชคดี, 2550; และศูนย์ส่งเสริมและพัฒนาพลังแผ่นดินเชิงคุณธรรม (ศูนย์คุณธรรม) สำนักงานบริหารและพัฒนาองค์ความรู้ (องค์การมหาชน), 2551) วิธีการที่ใช้ในการประเมินอภิमानส่วนมาก ประเมินหรือสังเคราะห์จากเอกสารรายงานต่างๆ รองลงมาคือ การสัมภาษณ์ทั้งเชิงลึกและทางโทรศัพท์จากกลุ่มตัวอย่าง นอกจากนี้ยังใช้เทคนิคสนทนากลุ่มอีกด้วย ส่วนการวิเคราะห์ข้อมูลมีทั้งเชิงคุณภาพ และเชิงปริมาณ ดังมีรายละเอียดในตารางที่ 2.11

ตารางที่ 2.11 สรุปงานวิจัยที่เกี่ยวข้องกับการประเมินอภิมานงานประเมิน

ชื่อผู้วิจัย	สิ่งที่ประเมิน	วิเคราะห์ข้อมูล
Widsor, Boyd, และ Orleans (1998)	ประเมินคุณภาพงานประเมินและสังเคราะห์ผลการประเมินโครงการเด็กสูบบุหรี่ในสตรีมีครรภ์ ใช้การสังเคราะห์งานวิจัยระหว่างปี ค.ศ.1986 ถึง 1998 จำนวน 23 เรื่อง	เชิงปริมาณ Meta-analysis
Bjornklide และ Bason (2000)	งานประเมินคุณภาพระดับอุดมศึกษาในประเทศเดนมาร์ก ระหว่างปีค.ศ. 1993 ถึง1997 โดยการใช้แบบสอบถามและการสัมภาษณ์เชิงลึกกับอธิการบดี และใช้กรณีศึกษาสำหรับ 6 คณะที่แตกต่างกัน	เชิงปริมาณและ เชิงคุณภาพ
Scott-little, Hamann, และ Jurs (2002)	ประเมินอภิมาน วิธีการและการสังเคราะห์ข้อค้นพบงานประเมินกิจกรรมหลังการเรียนในชั้นเรียนของสถานศึกษาในสหรัฐอเมริกา	เชิงคุณภาพ
Oxfam Australia (2006)	การศึกษาประเมินงานประเมินที่เสร็จสิ้นแล้ว โดย Oxfam ออสเตรเลียระหว่างมกราคม 2002 ถึง มิถุนายน 2006 โดย การทบทวนเอกสาร ทบทวนฐานข้อมูลการประเมิน 20 แห่ง สัมภาษณ์พนักงาน ประเมินเชิงลึกด้วยวิธีการและการเขียนรายงานเชิงประเมิน	ประเมินเชิงคุณภาพ
Madaus,และMcsherry (2007)	ประเมินประสิทธิผล โปรแกรม RCAB สำหรับปกป้องและป้องกันการกระทำผิดของเด็ก โดยการสัมภาษณ์โดยตรงและทางโทรศัพท์จาก พระสงฆ์และผู้นำศาสนา ผู้บริหารโรงเรียน ผู้ปกครอง และการสนทนากลุ่ม จากสมาชิกของหน่วยการปกครอง ครูในโรงเรียน ครูสอนศาสนา สมาชิกของทีม และอาสาสมัคร	เชิงคุณภาพและ เชิงปริมาณ
สมาคมคุศาสตรัฐศาสตร์แห่งประทศไทย (2547)	ประเมินอภิมานรายงานการประเมินคุณภาพสถานศึกษา การอาชีวศึกษา ประจำปีงบประมาณ 2545-2546 จำนวน 118 เล่ม	เชิงปริมาณและ เชิงคุณภาพ
วารุณี ลัภนโชคดี (2550)	ประเมินคุณภาพและสังเคราะห์ผลการประเมินคุณภาพของรายงานการประเมินตนเองและรายงานผลการประเมินคุณภาพภายนอกแรกของสถาบันอุดมศึกษาชนิดละ 200 เล่ม	เชิงปริมาณ
ศูนย์ส่งเสริมและพัฒนาพลังแผ่นดินเชิงคุณธรรม (ศูนย์คุณธรรม) สำนักงานบริหารและพัฒนาองค์ความรู้ (องค์การมหาชน) (2551)	ประเมินอภิมานเชิงสรุปรวม โครงการเร่งสร้างคุณลักษณะที่ดีของเด็กและเยาวชนไทย โดยการสังเคราะห์รายงานการประเมินโรงเรียนและรายงานการประเมินอภิมานของผู้ประเมินภายใน จำนวน 25 โรงเรียน	เชิงปริมาณและ เชิงคุณภาพ

6.3 การประเมินอภิमानจากผลงานการเรียน ได้แก่ งานวิจัยของ อมรรัตน์ พันธุ์งาม (2543) ทำศึกษาเพื่อพัฒนากระบวนการประเมินแฟ้มผลงานและการใช้ประโยชน์จากผลการประเมินแฟ้มผลงาน: การประยุกต์ใช้การประเมินอภิमान โดยทำการสังเคราะห์งานวิจัยที่เกี่ยวกับการประเมินแฟ้มผลงาน จากตำรา บทความทางวิชาการ และงานวิจัยด้วยการวิเคราะห์อภิमान พบว่า กลุ่มนักเรียนที่ใช้การประเมินแฟ้มผลงานมีผลสัมฤทธิ์ทางการเรียนสูงกว่ากลุ่มนักเรียนที่ไม่ใช้การประเมินแฟ้มผลงาน ข้อมูลเชิงคุณภาพจากการศึกษาภาคสนามโดยการสัมภาษณ์ครูในโรงเรียนประถมศึกษา และมัธยมศึกษา และประเมินผลแฟ้มผลงานด้วยการประเมินอภิमान สรุปได้ว่า รูปแบบการประเมินแฟ้มผลงานที่มีวัตถุประสงค์เพื่อประเมินความก้าวหน้ามีการดำเนินการประเมินครบทั้งการวางแผน การสะท้อนความคิด และการจัดทำแฟ้มผลงาน มีครูและนักเรียนเป็นผู้ประเมินและมีการอภิปรายและสะท้อนความคิดต่อผลการประเมิน ส่วนงานวิจัยของ ภัทราวรรณ ภิรมย์วงศ์ (2546) ทำการวิเคราะห์และประเมินอภิमान การประเมินผลการปฏิบัติงานของนักเรียน ในวิชาศิลปศึกษา ใช้การประเมินอภิमानประเมินวิธีการประเมินผลวิชาศิลปศึกษา ทำการศึกษา 2 วิธี วิธีแรกเป็นการศึกษาเชิงปริมาณ จากครูศิลปศึกษาในกรุงเทพฯ ด้วยแบบสำรวจ วิธีที่สอง เป็นการศึกษาระดับคุณภาพจากกรณีศึกษา 3 โรงเรียน ด้วยแบบสัมภาษณ์ และแบบสังเกตพฤติกรรม การปฏิบัติงานของครู และนำผลการวิจัยมาทำการประเมินอภิमानโดยใช้ผู้ทรงคุณวุฒิด้านการสอน ศิลปศึกษาด้วยการจัดประชุมสนทนากลุ่ม ทำการสังเคราะห์ผลการวิจัยโดยใช้การวิเคราะห์เนื้อหาเชิงบรรยาย ผลการวิเคราะห์เชิงปริมาณ พบว่า องค์ประกอบที่ใช้ในการประเมินและน้ำหนัก องค์ประกอบในการประเมินมี 4 องค์ประกอบ ได้แก่ กระบวนการปฏิบัติงาน ผลงาน พฤติกรรมของนักเรียน และความรู้พื้นฐานทางศิลปะ วิธีการประเมินผล ได้แก่ การประเมินจากผลงานการทดสอบ ภาคปฏิบัติ การสังเกต/พฤติกรรมนักเรียนของนักเรียน การมีส่วนร่วมในชั้นเรียน การทดสอบ ข้อเขียน สารของรายงานผล คือ พัฒนาการของผลงาน พฤติกรรมและความสนใจในการเรียน จุดเด่นและจุดด้อยของผลงาน โดยที่โรงเรียน A เป็นโรงเรียนที่มีรูปแบบที่ดีที่สุด ในด้านการจัด กิจกรรมการเรียนการสอนและการวัดและประเมินผลที่ถูกต้องตามหลักการประเมินผลการปฏิบัติงานมากที่สุด

ตอนที่ 4 แนวทางการเขียนรายงานการประเมิน

รายงานการประเมินที่ดี ควรประกอบด้วย ส่วนประกอบที่สำคัญของรายงาน ได้แก่ รายละเอียดของปกรายงาน คำนำ สารบัญ บทนำ วัตถุประสงค์การประเมิน ทฤษฎีและกรอบแนวทางการประเมิน ผลการประเมินตามมาตรฐานการประเมิน สรุปผลการประเมิน ข้อเสนอแนะ และภาคผนวก ดังนั้น จึงได้สังเคราะห์ส่วนประกอบของรายงานการประเมินที่มีมาตรฐาน คือ รายงานการประเมินของ UNDP (United Nations Development Programme) (2008), USAID

(United States Agency International Development) (2006), Victory University ในการประเมินงาน e-learning, รูปแบบมาตรฐานสำหรับการเขียนรายงานการประเมินของ DanChurchAid (2007), และรายงานประเมิน Ex-detainees Rehabilitation Program. (2002) และนำรายละเอียดดังกล่าวมาพิจารณาเปรียบเทียบกับรายงานการประเมินคุณภาพภายใน ตามที่สำนักงานคณะกรรมการการอุดมศึกษากำหนด ซึ่งจะเห็นได้ว่า รายงานการประเมินของสกอ. มีส่วนประกอบใกล้เคียงกับรายงานการประเมินของหลาย ๆ แห่ง แต่ ยกเว้นยังไม่ได้มีการกำหนดส่วนสำคัญดังต่อไปนี้ คือ กิตติกรรมประกาศ ข้อจำกัดของการประเมิน และบทเรียนที่ได้รับ แต่ในส่วนของคำย่อ และคำอธิบาย คำศัพท์เฉพาะ หากมีอธิบายไว้ในคู่มือการประเมินคุณภาพภายในแล้ว ทุกหน่วยงานเข้าใจตรงกัน ก็ไม่จำเป็นต้องระบุไว้ในรายงานอีก รายละเอียดสรุปดังตารางที่ 2.12

ตารางที่ 2.12 ส่วนประกอบหลักในรายงานการประเมิน

ส่วนประกอบ	UNDP	USAID	Victoria University*	Standard format for evaluation reports	Ex-detainees Rehabilitation Program. (2002)	รายงานประเมินของสกอ.
1. หน้าปก	✓	✓	✓		✓	✓
2. รายละเอียดโครงการประเมิน/ คำนำ		✓	✓		✓	✓
3. คำย่อ	✓	✓				
4. กิตติกรรมประกาศ		✓				
5. สารบัญ ตาราง รูปภาพ	✓	✓	✓		✓	✓
6. คำอธิบายคำศัพท์เฉพาะ		✓				
7. บทสรุปผู้บริหาร	✓	✓	✓	✓	✓	✓
8. บทนำ	✓	✓	✓	✓	✓	✓
9. วัตถุประสงค์การประเมิน	✓	✓		✓		✓
10. ขอบเขต และข้อจำกัดของการประเมิน	✓			✓		
11. การออกแบบและวิธีการประเมิน	✓	✓			✓	✓
12. ผลการประเมิน	✓	✓	✓	✓	✓	✓
13. สรุปผลการประเมิน	✓	✓	✓	✓	✓	✓
14. ข้อเสนอแนะ	✓	✓	✓	✓	✓	✓

ตารางที่ 2.12 (ต่อ)

ส่วนประกอบ	UNDP	USAID	Victoria University*	Standard format for evaluation reports	Ex-detainees Rehabilitation Program. (2002)	รายงานประเมินของ สกอ.
15. บทเรียนที่ได้รับ ทางเลือก	✓	✓		✓		
16. ภาคผนวก	✓	✓	✓	✓	✓	✓

* Lisa Emerson. (1995)

รายงานการประเมินเป็นผลผลิตหลักของกระบวนการประเมิน มีวัตถุประสงค์เพื่อหาผลการประเมินที่มีข้อมูลพื้นฐานที่โปร่งใส ตรวจสอบได้ สำหรับใช้ในการตัดสินใจเกี่ยวกับนโยบายและโครงการ เรียนรู้ ใช้ในการพัฒนาและปรับปรุงการเรียนรู้ ซึ่งคำอธิบายส่วนประกอบหลักของรายงานการประเมิน โดยรายงานต้องสามารถนำไปใช้งานง่ายและอ่านง่าย มีข้อเสนอแนะเชิงนโยบาย โดยต้องคำนึงถึงผู้ใช้รายงานเป็นสำคัญ มีความยาวไม่เกิน 60 หน้า (ไม่รวมภาคผนวก) โครงสร้างรายงานมีการอธิบายเนื้อหาในแต่ละส่วนชัดเจนสรุปได้ว่า ควรมีรายละเอียดดังต่อไปนี้

1. หน้าปก (Title page) แสดงหัวข้อที่ประเมิน ให้ข้อมูลที่ครอบคลุมบริบทพื้นฐาน การประเมินอย่างชัดเจน ได้แก่ วัน เวลา ที่ประเมิน สถานที่ประเมิน ทู่นในการดำเนินการ (ถ้ามี)

2. บทสรุปผู้บริหาร (Executive summary) เป็นส่วนที่สำคัญที่สุดของรายงาน ซึ่งมีรายละเอียดที่ผู้อ่านสามารถเข้าใจได้ง่าย เป็นการสรุปข้อมูลเกี่ยวกับวัตถุประสงค์การประเมิน ข้อค้นพบที่สำคัญ ข้อสรุป ข้อเสนอแนะที่เป็นประโยชน์ และบทเรียนที่ได้รับ โดยสรุปข้อจำกัด และความชัดเจนในงานประเมิน โดยสรุปรายละเอียดจากในเล่มรายงาน มีความยาวประมาณ 2-5 หน้า

3. บทนำ (Introduction) เป็นรายละเอียดเกี่ยวกับบริบท (Context) ที่ตั้งสถาบัน สรุปประวัติความเป็นมา ภูมิหลังของสิ่งที่ประเมินของสถาบันและผู้มีส่วนได้ส่วนเสีย วัตถุประสงค์ขอบเขตในภาพรวมของการประเมิน และคำนึงถึงผู้ใช้ผลการประเมิน สาเหตุของปัญหา รายละเอียดของกิจกรรมที่ดำเนินการ ผลผลิตหลักและค่าใช้จ่ายโดยสรุป ตลอดจนนโยบายและการพัฒนาการประเมิน รวมถึงข้อตกลงเบื้องต้นเกี่ยวกับ ปัจจัยภายนอกที่มีส่วนในการวางแผน ซึ่งผู้ประเมินจะสรุปข้อค้นพบและข้อสรุปเบื้องต้นในการประเมิน สภาพเศรษฐกิจ การเมือง ชุมชน และประเทศชาติ หากมีรายละเอียดเกี่ยวกับบริบทเฉพาะของสถานที่ประเมิน จะช่วยให้ผู้อ่านเข้าใจสภาพแวดล้อมที่แท้จริงของสถานที่ประเมิน นอกจากนี้อาจมีแผนที่บริเวณที่ประเมินและสถานที่ใกล้เคียง รวมถึงรูปภาพประกอบ มีการสรุปสั้นๆ ถึงวิธีการประเมินที่ประยุกต์ใช้เฉพาะ

กลุ่มที่รับการประเมินและผู้มีส่วนได้ส่วนเสียอื่นที่เกี่ยวข้อง ในส่วนนำนี้จะอธิบายถึง สภาพปัญหาของสิ่งที่ประเมิน ทฤษฎีที่ใช้ในการประเมิน รวมถึงการออกแบบการประเมินที่สอดคล้องกับสภาพและบริบทของสิ่งที่ประเมิน และกรอบแนวทางในการวิพากษ์รายงานการประเมิน เสนอแนะแนวทางแก่ผู้อ่านรายงานการประเมินเพื่อนำไปปรับปรุงการประเมินครั้งต่อไป

4. การออกแบบและวิธีการประเมิน (Design and evaluation methodology)

เป็นการบรรยายลักษณะรายงานการประเมิน ที่น่าเชื่อถือ การแสดงหลักฐานสนับสนุนคุณภาพการประเมิน ขึ้นอยู่กับการออกแบบการประเมิน วิธีการเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล โดยอาจต้องใช้นักประเมินมืออาชีพ และนักวิจัยที่สามารถอธิบายวิธีการรวบรวมข้อมูลที่ชัดเจน

1) การออกแบบการประเมิน ได้แก่ การประเมินกึ่งทดลอง (quasi-experimental) วิธีผสมผสาน (mixed methods), การรวบรวมข้อมูลระยะยาว (longitudinal data), การเปรียบเทียบข้อมูลอนุกรมเวลา (time-series comparison) การทดลองแบบสุ่ม (randomized trial)

2) วิธีการเก็บรวบรวมข้อมูล ได้แก่ การเลือกกลุ่มตัวอย่าง เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ชนิดของข้อมูล เทคนิคการวิเคราะห์ข้อมูล เขียนอธิบายสรุปสั้นๆ ซึ่งข้อมูลเหล่านี้ อาจใส่ไว้ในภาคผนวก รวมถึงแบบสอบถาม แบบสังเกต รายละเอียดการสุ่มตัวอย่าง Printout ผลการวิเคราะห์ข้อมูล และเอกสารอื่นๆ

3) แหล่งข้อมูล ควรมีการเปรียบเทียบจากหลายแหล่ง หรือเปรียบเทียบการประเมินก่อน และหลังการประเมิน ข้อมูลที่ใช้ในการวิเคราะห์ข้อมูลต้องมีความพอเพียง

4) ช่วงเวลาและทรัพยากร จำเป็นต้องมีการพิจารณาถึงผลกระทบต่อข้อค้นพบและข้อสรุปในการประเมิน รวมถึงข้อจำกัดในการประเมิน

5. ผลการประเมิน (Finding) ข้อค้นพบหรือผลการประเมินเป็นข้อมูลเชิงประจักษ์ที่ผู้ประเมินเสนอหลักฐานที่เกี่ยวข้องกับคำถามในการประเมิน ข้อค้นพบจะนำเสนอตามวัตถุประสงค์ของทีมประเมิน ผู้อ่านจะทราบได้จากความคิดเห็นเกี่ยวกับจุดแข็งและจุดอ่อนของข้อสรุปในการประเมิน ซึ่งจะนำเสนอในหลายช่องทาง ประกอบด้วย 1) กระบวนการค้นหาคำตอบเกี่ยวกับการจัดการและการดำเนินการโครงการ 2) ผลลัพธ์และผลผลิตจากข้อค้นพบ เกี่ยวกับการบรรลุเป้าหมายของโครงการ 3) การนำเสนอข้อค้นพบ ด้วยตาราง กราฟ รูปภาพ 4) ข้อค้นพบที่เกิดขึ้นจริงจากเอกสาร การสังเกต การสัมภาษณ์ และการรวบรวมข้อมูลอื่นๆ 5) การวิเคราะห์ข้อมูล โดยเริ่มตั้งแต่การรวบรวมข้อมูลดิบ คำนวณหาค่าความถี่ ร้อยละ เปรียบเทียบข้อมูลแต่ละกลุ่ม เปรียบเทียบกับมาตรฐาน เปรียบเทียบกับเป้าหมาย อธิบายลักษณะข้อมูลจากการวิเคราะห์ ได้แก่ correlation tests, cross-tabulations, regression analysis, factor analysis, การ

วิเคราะห์เชิงคุณภาพ (qualitative analysis) การตัดสินประเมินขึ้นอยู่กับความคิดเห็นที่สอดคล้องกัน ประสิทธิภาพ ประสิทธิผลและความยั่งยืน

6. ข้อสรุป (Conclusion) เป็นการสรุปเพื่อตัดสินผลการประเมิน ดีความจากข้อค้นพบ เป็นการตัดสินความสำคัญ/ดีที่ที่สุด หรือคุณค่าของกิจกรรม/โครงการที่ประเมิน จากข้อมูลเชิงประจักษ์ (empirical finding) เกี่ยวกับบริบท ปัญหา การจัดการและการปฏิบัติการประเมิน ซึ่งการสรุปผลเป็นการตัดสินผลการประเมินของทีมผู้ประเมิน ใช้ในการเขียนรายงานการประเมิน ข้อค้นพบจากการประเมินและผลการวิเคราะห์ข้อมูล เพื่อให้ข้อเสนอแนะต่อไป สำหรับความน่าเชื่อถือของข้อสรุปมาจากหลักฐานอ้างอิงและความเชี่ยวชาญของทีมผู้ประเมินเป็นสำคัญ

7. ข้อเสนอแนะ (Recommendations) ข้อเสนอแนะจากการประเมินมาจากข้อค้นพบ ข้อสรุปจากการประเมินที่ชี้ให้เห็นว่ามีความจำเป็นต้องเสนอแนะ ตัวบ่งชี้การปฏิบัติการประเมินของผู้ประเมิน รวมถึง แหล่งทรัพยากร การเงิน การออกแบบ การดำเนินการและการติดตามการประเมิน รวมถึงความต้องการและความสนใจในการประเมิน มีข้อมูลสนับสนุนเหตุผลในการวิเคราะห์ผลการประเมิน อีกทั้งข้อเสนอแนะต้องเป็นประโยชน์ต่อผู้มีส่วนได้ส่วนเสียหลักในการประเมิน

8. บทเรียนที่ได้รับจากการประเมิน (Lessons Learned) เป็นผลกระทบจากการประเมิน ที่เสนอโอกาสการเรียนรู้ อาจจะสรุปอ้างอิงได้เหนือกว่าผลการประเมินจากโครงการประเมิน มาจากประสบการณ์อันยาวนานและมุมมองของผู้ประเมิน มากกว่าข้อค้นพบระหว่างการประเมิน เพื่อเสนอแนะแนวปฏิบัติที่ดีตามบริบทของสถาบัน

9. ภาคผนวก (The Annex) เป็นการรวบรวมเอกสารสนับสนุนรายงานการประเมินที่สำคัญ แต่ไม่ได้ใส่ไว้ในรายงาน เพื่อช่วยให้ผู้อ่านสามารถสืบค้นได้ ได้แก่ ขอบเขตของการประเมิน รายละเอียดเกี่ยวกับวิธีการประเมิน การเก็บรวบรวมข้อมูล เครื่องมือ รายละเอียดผู้ให้สัมภาษณ์ รายละเอียดการวิเคราะห์ข้อมูลเชิงลึก และเอกสารอ้างอิงอื่นๆ เป็นต้น รายละเอียดของวิธีการที่ประยุกต์ใช้ (รวมถึงการออกแบบการวิจัย การสุ่มตัวอย่าง เครื่องมือในการเก็บรวบรวมข้อมูล) กระบวนการในการวิเคราะห์ข้อมูล การอธิบายถึงข้อจำกัด และจุดแข็งในการเลือกวิธีการในการประเมิน

รายงานการประเมินที่ดี ต้องมีเกณฑ์คุณภาพ คือ อรรถประโยชน์ (utility) ความน่าเชื่อถือ (credibility) และความสอดคล้อง เหมาะสม (relevance/appropriateness)

1. อรรถประโยชน์ (utility) เป็นรายงานการประเมินที่เป็นประโยชน์ เมื่อ 1) ข้อมูลในรายงานมีความสมบูรณ์ในบริบทการประเมิน คุณค่าการประเมินมาจาก ผู้มีส่วนได้ส่วนเสียที่

เกี่ยวข้อง ความน่าเชื่อถือของนักประเมินและสถาบัน แนวทางของนักประเมินกับสถาบัน ในการแปล ความ งบประมาณ เวลา ความเกี่ยวพันและเป็นแนวทางเดียวกับชาติ 2) การนำเสนอกระบวนการ และข้อค้นพบจากการประเมินที่สมบูรณ์ จัดหาสารสนเทศที่จำเป็นสำหรับการตัดสินใจและลง ข้อสรุปในแต่ละส่วนของรายงาน 3) มีข้อเสนอแนะที่ชัดเจนและปฏิบัติได้ และ 4) สารสนเทศตาม แผนที่คาดหวังด้วยการประเมินตามที่ต้องการของผู้มีส่วนได้ส่วนเสีย

รายงานการประเมินที่ดี ควรประกอบด้วยตัวบ่งชี้ ดังต่อไปนี้ UNDP (United Nations Development Programme) (2008) คือ

1) หน้าปก แสดงหัวข้อที่ประเมิน ให้สารสนเทศที่ครอบคลุมบริบทพื้นฐานการ ประเมินอย่างชัดเจน

2) การประเมินมีการบรรยายถึงบุคคลที่เกี่ยวข้อง ที่มีบทบาทในการประเมิน

3) สารสนเทศในรายงานมีความสมบูรณ์ โครงสร้างครบถ้วน นำเสนอชัดเจน ได้แก่ วัตถุประสงค์ในการประเมิน สิ่งที่ประเมิน วิธีการออกแบบการประเมิน หลักฐานในการประเมิน การลงสรุป ข้อเสนอแนะ และบทเรียนที่ได้รับจากการประเมิน

4) รายงานการประเมินจะต้องมีความชัดเจน อ่านเข้าใจง่าย ประกอบด้วย รูปภาพ กราฟ ตารางและตัวอย่างที่ง่ายต่อความเข้าใจ กระชับ มีภาคผนวก รวมถึง ระยะเวลาในการ ประเมิน การดำเนินการและงบประมาณ

5) บทสรุปผู้บริหารควรสรุปสั้น และมีสารสนเทศที่สำคัญ ได้แก่ ข้อค้นพบหลัก ข้อสรุป และข้อเสนอแนะที่สำคัญ ต่อการตัดสินใจ

6) ข้อเสนอแนะที่สอดคล้องกับ ผลการประเมินและข้อสรุป รวมถึงสอดคล้องกับ วัตถุประสงค์ของการประเมินและการตัดสินใจโดยอิงจากการประเมิน โดยต้องมีความชัดเจนและ สามารถนำไปปฏิบัติได้จริงทันเหตุการณ์

2. ความน่าเชื่อถือ (credibility) เป็นรายงานการประเมินที่น่าเชื่อถือ เมื่อมีการใช้ กระบวนการและเครื่องมือที่มีความเป็นปรนัย ความตรง และความเที่ยง 1)นักประเมินมีสมรรถนะ ในวิชาชีพ และตรงกับผู้มีส่วนได้ส่วนเสียต้องการ 2)มีความถูกต้อง และความตรง (เนื้อหาความ สิ่ง ที่ประเมิน และปัจจัยแวดล้อม เครื่องมือ สารสนเทศ ความตรงภายนอกหรือเชื่อมโยงกับการ พัฒนาข้อค้นพบ 3) มีความเที่ยงหรือสอดคล้องกับสารสนเทศ และ4)เป็นพื้นฐานในการตัดสินใจ ที่โปร่งใส อิงข้อตกลงอย่างเป็นทางการ ประกอบด้วยตัวบ่งชี้ ดังนี้

1) สิ่งที่จะประเมิน ได้แก่ เป้าหมายและวัตถุประสงค์ บริบท การดำเนินการ และ ประโยชน์ที่คาดหวังของการประเมิน มีความชัดเจน และถูกต้อง

2) รายงานการประเมินมีการอธิบายถึงขอบเขตในการประเมินได้อย่างชัดเจน และบอกถึงข้อจำกัดของการประเมิน

3) วิธีการประเมินที่สามารถอธิบายถึงความตรงและความเที่ยงในการประเมิน ข้อมูลในการตอบคำถามการประเมินและแหล่งข้อมูล กระบวนการสุ่มตัวอย่างที่เป็นตัวแทนในการรวบรวมข้อมูล กระบวนการประเมินที่ความถูกต้องและสารสนเทศที่รวบรวมมาได้อย่างน่าเชื่อถือ การตัดสินใจ การตีความข้อค้นพบสรุปตามตัวบ่งชี้หรือมีระดับนัยสำคัญทางสถิติ กระบวนการประเมินวิเคราะห์จากข้อมูลทั้งเชิงคุณภาพและเชิงปริมาณ มีความเสมอภาคในการออกแบบตามเงื่อนไขของสารสนเทศที่แตกต่างกัน การพัฒนานวัตกรรมและการปกป้องสิทธิมนุษยชนในการประเมิน และขอบเขตการประเมิน

4) ข้อค้นพบจากการประเมินตามคำถามการประเมิน รวมถึงประสิทธิภาพในการประเมิน ปัจจัยภายนอกที่มีอิทธิพลต่อผลลัพธ์การประเมิน

5) ข้อสรุปจากการตัดสินใจคุณค่าของผู้ประเมินที่สอดคล้อง มีประสิทธิภาพ ประสิทธิภาพ และความยั่งยืน โดยอิงหลักฐานและการวิเคราะห์ผลการประเมิน

6) ภาคผนวกที่สมบูรณ์และสอดคล้องกัน ได้แก่ แผน/กำหนดการประเมิน รายละเอียดของข้อมูลและการวิเคราะห์ เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล

3. ความสอดคล้อง เหมาะสม และเพิ่มคุณค่า (relevance, appropriateness and added-value) รายงานสอดคล้อง เหมาะสม มีคุณค่าเมื่อมีสารสนเทศเหมาะสมหรือสารสนเทศที่จำเป็น เหมาะสมกับเป้าหมายของสถาบัน การทำการประเมินที่เป็นแนวทางเดียวกับระบบการประเมินของชาติ 1) วัตถุประสงค์และกระตุ้นให้ผู้ให้ข้อมูลมีความชัดเจน 2) มีแนวทางการประเมินที่สอดคล้องกับชาติและคำนึงถึงสิทธิมนุษยชน 3) เหมาะสมกับแผนกลยุทธ์ และ 4) มีความรู้หรือสิทธิมากขึ้นสำหรับการพัฒนา (ความเท่าเทียม ความสามารถ ความร่วมมือ และอื่นๆ) ประกอบด้วยตัวบ่งชี้ดังต่อไปนี้

- 1) การอธิบายวัตถุประสงค์และบริบทของการประเมิน
- 2) การรายงานสรุปการประเมินที่คำนึงถึง ความเท่าเทียมกันระหว่างเพศ และพิจารณาสิทธิมนุษยชนของผู้รับการประเมินในการออกแบบและความแตกต่างของสารสนเทศในการประเมินภายในขอบเขตของการประเมินและวิธีการประเมินที่ใช้
- 3) รายงานนำเสนอสารสนเทศที่มีความสัมพันธ์กับการประเมินอื่น และบ่งชี้ถึงการเพิ่มคุณค่าของสารสนเทศ

4. กระบวนการและความรับผิดชอบที่ตรวจสอบได้ (procedures and accountabilities)

ความรับผิดชอบขั้นต้น ในการเตรียมรายงานการประเมินร่วมกับที่ปรึกษาการประเมิน หัวหน้าทีมการประเมิน และคณะกรรมการประเมิน ซึ่งผู้มีส่วนได้ส่วนเสียหลักจะสามารถเข้ามามีส่วนร่วมในการทบทวนร่างรายงานการประเมินเพื่อตรวจสอบความสอดคล้องของข้อเท็จจริง ข้อผิดพลาด หรือข้อยกเว้น ที่มีผลต่อการตีความข้อค้นพบจากการประเมินที่อาจทำให้เกิดการพิจารณาที่ไม่ถูกต้อง ผู้ประเมินจะสามารถขอรับการเปลี่ยนแปลงเพื่อแก้ไขให้ถูกต้อง แต่ทั้งนี้ยึดหลักการความเป็นอิสระของผู้ประเมินที่จะสามารถสรุปผลการประเมินจากข้อค้นพบด้วยตนเอง

งานวิจัยที่เกี่ยวข้อง

Griffin และ Miyoshi (2009) ศูนย์การวิจัยทางการประเมิน มาตรฐานและการทดสอบทางการศึกษาแห่งชาติ (CRESST) ได้รับการประเมินภายนอกในรอบ 3 ปี จากโครงการประเมินการเรียนรู้แบบอัจฉริยะ โมเดลการพัฒนาโรงเรียนอิงความสามารถพัฒนามาจากการทำงานและปรัชญาของ Leonard Bernstein นักแต่งเพลงคนล่าสุด รายงานฉบับนี้เป็นรายงาน 3 ปี จากข้อค้นพบจากการประเมินการมีส่วนร่วมในการเรียนรู้แบบอัจฉริยะในช่วงปี 2003-2004 วัตถุประสงค์เพื่อให้สารสนเทศเกี่ยวกับการดำเนินการและผลกระทบของโครงการต่อโรงเรียนที่เข้าร่วมโครงการ ข้อค้นพบนี้อยู่ภายในบริบทของข้อค้นพบในภาพรวมจาก 3 ปีที่เข้าร่วม ใช้วิธีการรวบรวมข้อมูลการประเมินด้วยวิธีการเชิงปริมาณที่หลากหลายและวิธีเชิงคุณภาพ ข้อค้นพบเสนอแนะว่า โครงการการเรียนรู้แบบอัจฉริยะเป็นเครื่องมือที่มีประโยชน์ต่อครู ด้วยการสอนที่หลากหลายวิธี สอดคล้องตามความต้องการของมลรัฐ เนื้อหาที่สอน และกลุ่มนักเรียน ความพึงพอใจของครูต่อองค์ประกอบของโครงการในการพัฒนาวิชาชีพ อยู่ในระดับสูง แม้ว่าการประเมินจะเลือกเฉพาะเพื่อสนับสนุนความต้องการ ข้อเสนอแนะและอภิปราย จากการประเมินทั้ง 3 ปี

Lin, Lin, และ Chern (2009) โครงการพัฒนาการผลิตยาในระดับโลก ถูกพัฒนาขึ้นเพื่อพัฒนาการผลิตยาและทำการตลาดในเอเชีย ซึ่งมีการจัดทำรายงานผลการประเมินทางเศรษฐกิจ (PER) เริ่มจากไต้หวัน จากประเทศที่มีความร่วมมือกันทางด้านเศรษฐกิจ (APEC) และดำเนินการในเปรู ในปี ค.ศ. 2008 ในสถานการณ์ต่างกัน เพื่อเสนอแนะผู้นำ APEC การจัดทำรายงาน PER เพื่อเป็นการเผยแพร่สารสนเทศเกี่ยวกับผลิตภัณฑ์ยาชนิดใหม่ ในประเทศต่างๆ จึงมีการแลกเปลี่ยนรายงานกัน 17-21 ประเทศทั่วโลก โดยบริษัทต่างๆจะเห็นรายงานและให้ข้อเสนอต่างๆ เพื่อเป็นการพัฒนาคุณภาพผลิตภัณฑ์ยาให้ดีขึ้น

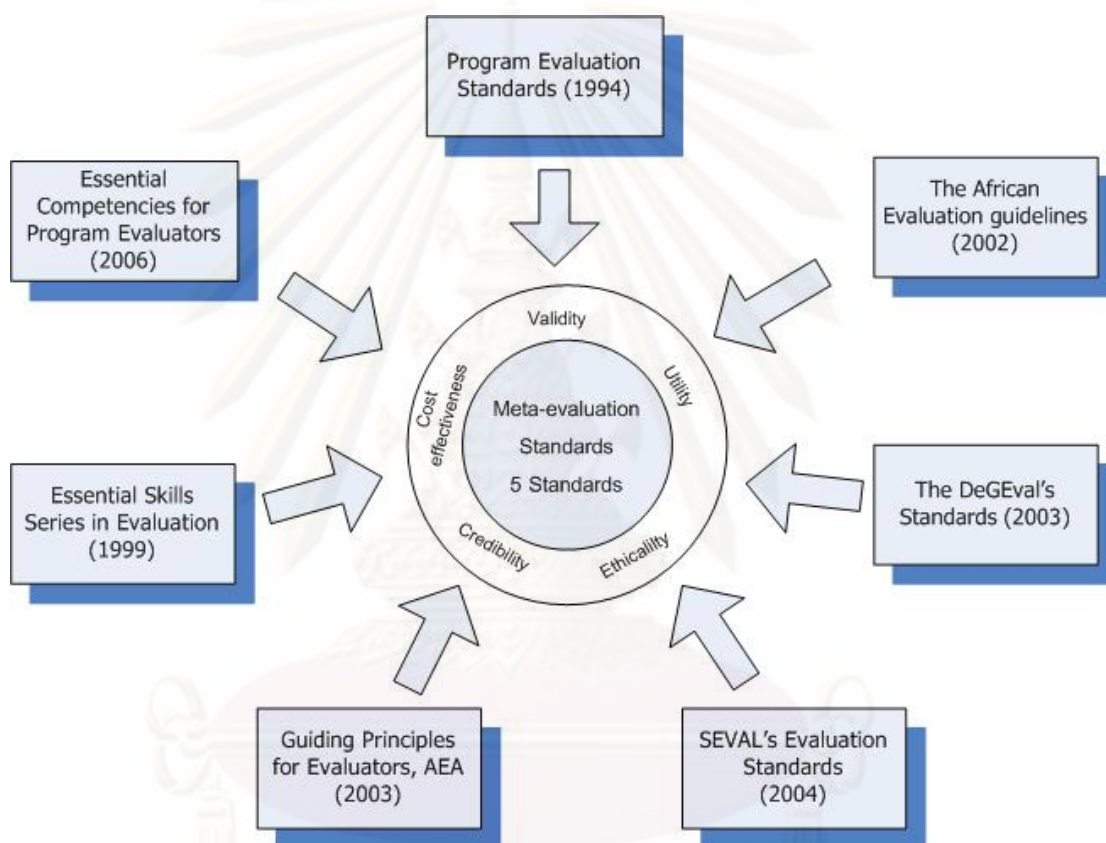
ตอนที่ 5 กรอบแนวคิดในการวิจัย

การพัฒนามาตรฐานการประเมินอภิमान ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับมาตรฐานการประเมิน การประเมินอภิमान และนำแนวคิดเทคนิค กระบวนการประเมินอภิमानไปใช้ในการประเมินรายงานการประเมินคุณภาพภายใน ระดับอุดมศึกษา

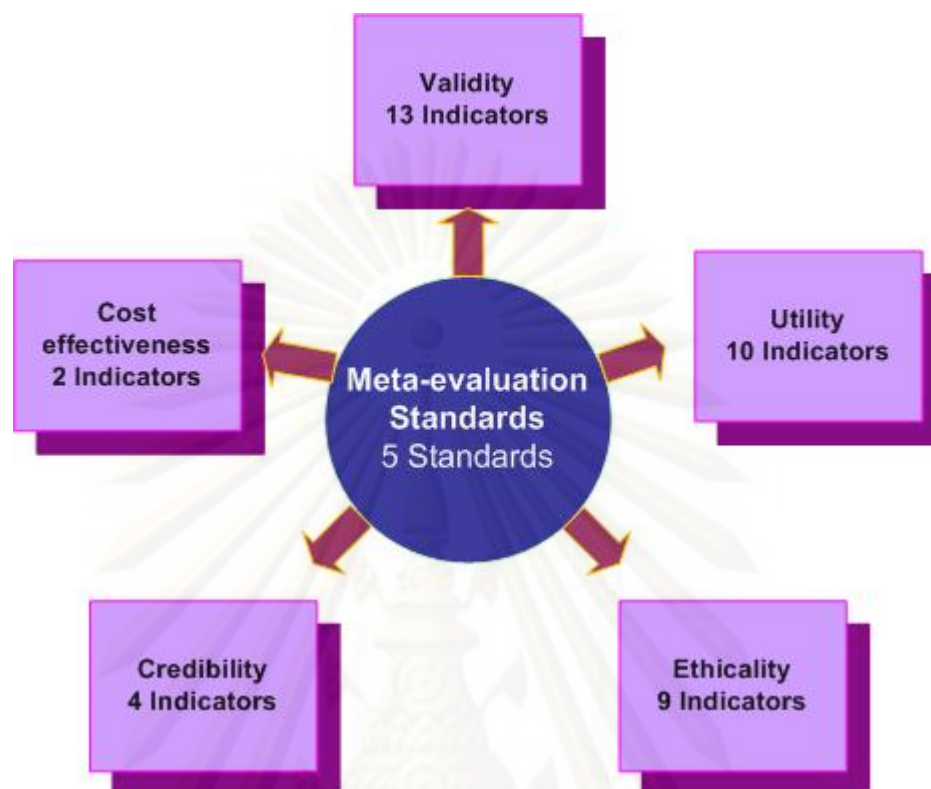
ซึ่งกรอบแนวคิดในการวิจัยในครั้งนี้ เพื่อพัฒนามาตรฐานการประเมินอภิमानของประเทศไทย และแนวคิดในการประยุกต์ใช้มาตรฐานการประเมินอภิमान ในการพัฒนาเครื่องมือในการประเมินอภิमान ได้แก่ แบบประเมินอภิमान คู่มือการประเมินอภิमान และหลักสูตรฝึกอบรมพัฒนา นักประเมินอภิमान โดยกำหนดวิธีการ ขั้นตอนในการประเมินอภิमान เมื่อนำนักประเมินมาพัฒนาตามหลักสูตรการพัฒนานักประเมินอภิमान จะทำให้มีทักษะในการประเมินอภิमान โดยสามารถทำการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษาได้ ตามมาตรฐานการประเมินอภิमान

การพัฒนามาตรฐานการประเมินอภิमानสำหรับประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษาในประเทศไทย ดังนั้นกรอบแนวคิดในการวิจัยครั้งนี้ จึงเกี่ยวข้องกับการพัฒนามาตรฐาน และเกณฑ์การประเมินอภิमान เพื่อนำไปประยุกต์ใช้ในการประเมินอภิमान ซึ่งประกอบด้วย การพัฒนามาตรฐานการประเมินอภิमान เพื่อให้เหมาะสมกับบริบทประเทศไทย ในกรอบมาตรฐานการประเมิน 5 มาตรฐานของ Davidson, 2005 ได้แก่ 1) ความตรงของการประเมิน (validity) 2) อรรถประโยชน์ (utility) 3) จริยธรรมในการประเมิน (ethicality) 4) ความน่าเชื่อถือของนักประเมิน (credibility) และ 5) ประสิทธิภาพต้นทุน (cost-effectiveness) โดยทำการสังเคราะห์จากมาตรฐานการประเมินที่เกี่ยวข้องใน 7 แนวทางการประเมิน คือ 1) Program Evaluation Standards ตามแนวทางของ the Joint Committee on Standard for Educational Evaluation (1994) ประกอบด้วย 4 มาตรฐานหลัก 30 มาตรฐานย่อย 2) The African Evaluation guidelines 2002 (AEG) ที่พัฒนาขึ้นโดยสมาคมการประเมินของประเทศแอฟริกา (African Evaluation Association) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 30 มาตรฐานย่อย 3) มาตรฐานการประเมินในบริบทของประเทศสวิตเซอร์แลนด์ เรียกว่า SEVAL's Evaluation Standards พัฒนาโดยสมาคมการประเมินของประเทศสวิตเซอร์แลนด์ (Swiss Evaluation Society) กำหนดมาตรฐานการประเมินหลัก 4 มาตรฐาน และ 28 มาตรฐานย่อย (Laubli Loud, 2004) 4) The DeGEval's Standards (2003) ที่พัฒนาโดยสมาคมการประเมินของประเทศเยอรมัน (German Evaluation Society) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 25 มาตรฐานย่อย 5) หลักการและแนวทางการประเมินสำหรับผู้ประเมิน Guiding Principles for

Evaluators ของ สมาคมประเมินแห่งสหรัฐอเมริกา (American Evaluation Association [AEA] ,2003) 6) ทักษะที่จำเป็นในการประเมิน Essential Skills Series in Evaluation ของสมาคมประเมินแห่งแคนาดา (Canadian Evaluation Society,1999) และ7) สมรรถนะที่จำเป็นสำหรับผู้ประเมินโครงการ Essential Competencies for Program Evaluators (Ghere & et al., 2006) ประกอบด้วย 38 ตัวบ่งชี้ ดังภาพที่ 2.3 และ 2.4



ภาพที่ 2.3 กรอบแนวทางในการพัฒนามาตรฐานการประเมินอภิมาน



ภาพที่ 2.4 กรอบมาตรฐานการประเมินอภิมานที่ตั้งเคราะห์ขึ้น

การประยุกต์ใช้มาตรฐานการประเมินอภิมาน และตรวจสอบคุณภาพของมาตรฐานการประเมินอภิมาน การดำเนินการแบ่งออกเป็น 3 ประเด็นหลัก คือ

1. การตรวจสอบคุณภาพมาตรฐานการประเมินอภิมาน โดยการนำมาตรฐานการประเมินอภิมาน และตัวบ่งชี้ ที่พัฒนาขึ้น ไปทำการสังเคราะห์เชิงทฤษฎีเทียบกับมาตรฐานการประเมินต่างๆ ดังที่กล่าวมาแล้ว และการประเมินอภิมานของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) เพื่อพิจารณาประเด็นความครอบคลุมและความเหมาะสมของตัวบ่งชี้ในการประเมินอภิมาน และหลังจากนำมาตรฐานการประเมินอภิมานไปประยุกต์ใช้กับข้อมูลเชิงประจักษ์ และทำการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) เพื่อพิจารณาความสอดคล้องของมาตรฐานและตัวบ่งชี้

2. พัฒนาเครื่องมือสำหรับประเมินอภิมาน ได้แก่ แบบประเมินอภิมาน คู่มือการประเมินอภิมาน และหลักสูตรฝึกอบรมนักประเมินอภิมาน เพื่อพัฒนานักประเมินอภิมานตามหลักสูตรฝึกอบรมพัฒนานักประเมินอภิมาน ทำการฝึกอบรมรวม 2 รุ่น โดยให้นักประเมินอภิมานทำการประเมินอภิมานรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา และนำผลการประเมินอภิมานมาเปรียบเทียบหาความสอดคล้องกัน เพื่อหาคุณภาพด้านความ

เที่ยง (reliability) และนำรายงานผลการประเมินคุณภาพภายในให้ผู้ทรงคุณวุฒิด้านการประกันคุณภาพเป็นผู้บ่งชี้คุณภาพ ด้วยเทคนิคกลุ่มรู้จัก (known group method) เพื่อพิจารณาหาคุณภาพด้านความตรงเชิงโครงสร้าง (construct validity) โดยเปรียบเทียบกับผลการประเมินคุณภาพโดยนักประเมินอภิมานที่พัฒนาขึ้น

3. การประยุกต์ใช้มาตรฐานการประเมินอภิมานที่พัฒนาขึ้น โดยออกแบบขั้นตอนวิธีการในการประเมินอภิมาน คู่มือการประเมินอภิมาน รายละเอียดดังต่อไปนี้

3.1 กำหนดการออกแบบ ขั้นตอน วิธีการในการประเมินอภิมาน ที่สอดคล้องกับมาตรฐานการประเมินอภิมาน โดยการเขียนคู่มือการประเมินอภิมาน

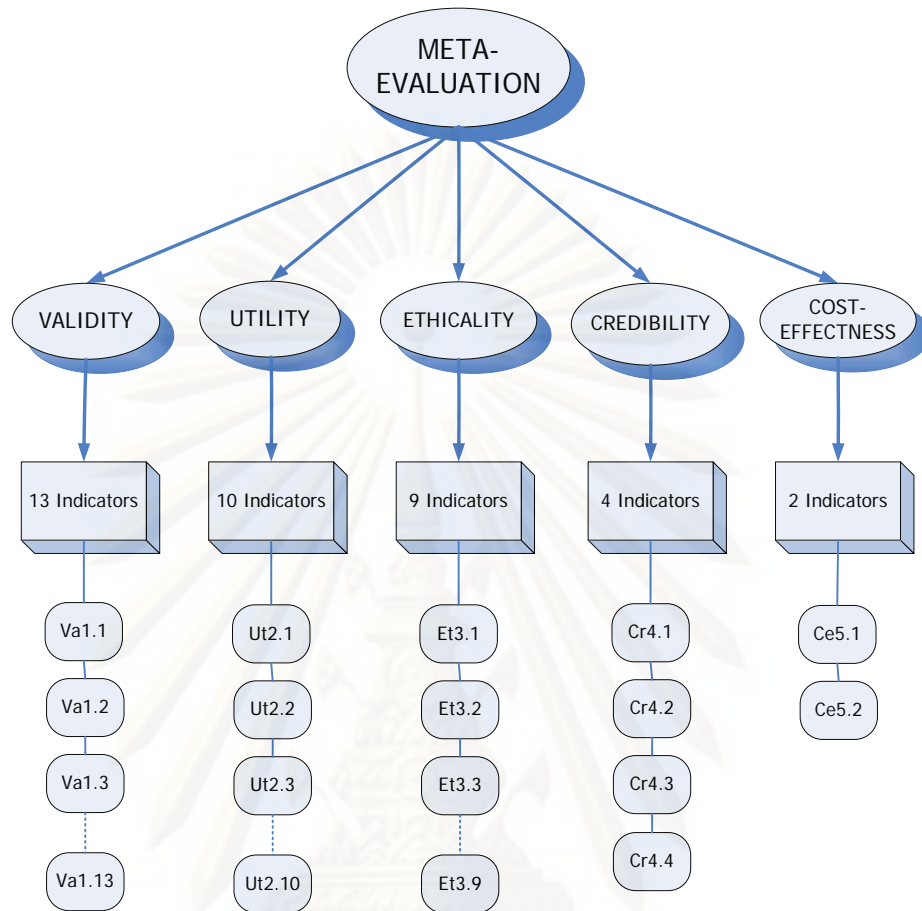
3.2 พัฒนาหลักสูตรในการฝึกอบรมนักประเมินอภิมาน ระดับอุดมศึกษาให้สอดคล้องกับมาตรฐาน และเกณฑ์การประเมินอภิมานที่พัฒนาขึ้น

3.3 ให้นักประเมินอภิมานที่ผ่านการอบรม ดำเนินการประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา

3.4 นำผลการประเมินอภิมานตามมาตรฐานการประเมินอภิมาน มาพิจารณาคุณภาพของรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาและนำผลมาวิเคราะห์เขียนข้อเสนอแนะต่อหน่วยงานที่เกี่ยวข้อง

ศูนย์วิทยทรัพยากร

จุฬาลงกรณ์มหาวิทยาลัย



ภาพที่ 2.5 กรอบแนวคิดในการพัฒนามาตรฐานการประเมินนิมิต

ศูนย์วิทยทรัพยากร

จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อพัฒนามาตรฐานการประเมินอภิมาน สำหรับประเมินรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาในประเทศไทย 2) เพื่อตรวจสอบคุณภาพของมาตรฐานการประเมินอภิมาน ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา 3) เพื่อประยุกต์ใช้มาตรฐานการประเมินอภิมานในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา ระดับอุดมศึกษา

สำหรับมาตรฐานการประเมินอภิมานที่พัฒนาขึ้นนำมาประยุกต์ใช้ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาเป็นการประเมินอภิมานแบบสรุปรวม (summative meta-evaluation) ประเมินโดยนักประเมินอภิมานภายนอก (external meta-evaluators) ที่ผ่านการอบรม เพื่อตรวจสอบคุณภาพของรายงานผลการประเมินคุณภาพภายใน ตรวจสอบความสอดคล้องของผลการประเมินอภิมานระหว่างนักประเมินอภิมาน และพิจารณาคุณภาพของรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษาตามมาตรฐานการประเมินอภิมาน ซึ่งมีรายละเอียดวิธีดำเนินการวิจัย แบ่งออกเป็น 5 หัวข้อ คือ 1. ขั้นตอนการดำเนินการวิจัย 2. แหล่งข้อมูลสำหรับการวิจัย 3. เครื่องมือและการพัฒนาเครื่องมือในการวิจัย 4. การเก็บรวบรวมข้อมูล และ 5. การวิเคราะห์ข้อมูล ดังมีรายละเอียดต่อไปนี้

1. ขั้นตอนการดำเนินการวิจัย

การวิจัยในครั้งนี้เป็นการวิจัยที่มุ่งประเมินอภิมาน โดยมีกรอบการดำเนินการดังต่อไปนี้ คือ

1. การศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องเพื่อพัฒนามาตรฐานการประเมินอภิมานของไทย จากมาตรฐานการประเมิน program evaluation standards และ Guiding Principles for Evaluators รวมถึงมาตรฐานการประเมินที่นำไปประยุกต์ใช้ในประเทศต่างๆ เพื่อพัฒนาเป็นแนวทางการพัฒนามาตรฐานการประเมินอภิมานในครั้งนี้

2. พิจารณา มาตรฐานการประเมินอภิมานที่พัฒนาขึ้น โดยตรวจสอบผลการสังเคราะห์มาตรฐานการประเมินอภิมาน ให้สอดคล้องกับการประเมินรายงานผลการประเมินคุณภาพภายใน และคู่มือการประเมินคุณภาพภายในสถานศึกษา ระดับอุดมศึกษา

3. ดำเนินการพัฒนาเครื่องมือที่ใช้ในการประเมินอภิमान คือ แบบประเมินอภิमान หลักสูตรฝึกอบรมนักประเมินอภิमान และคู่มือการประเมินอภิमान รายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา

4. ให้ผู้เชี่ยวชาญจำนวน 5 ท่าน ที่มีคุณสมบัติเป็นผู้มีความรู้ ความสามารถ เกี่ยวกับการประเมินคุณภาพการศึกษา ในระดับอุดมศึกษา และ/หรือการวัดผลการศึกษา พิจารณา ตรวจสอบคุณภาพเครื่องมือที่ใช้ในการประเมินอภิमान ด้านความตรงตามเนื้อหา (content validity) และดำเนินการปรับปรุงเครื่องมือตามข้อเสนอแนะของผู้เชี่ยวชาญให้เหมาะสมในการประเมิน

5. ทำการศึกษาหลักสูตรในการฝึกอบรมนักประเมินอภิमान การศึกษาขั้นพื้นฐาน จากศูนย์เครือข่าย สมศ. จุฬาลงกรณ์มหาวิทยาลัย และดำเนินการออกแบบเพื่อพัฒนาหลักสูตรในการฝึกอบรมนักประเมินอภิमानภายนอก ในระดับอุดมศึกษาเพื่อประเมินคุณภาพรายงานการประเมินคุณภาพของสถาบันอุดมศึกษา ตามมาตรฐานการประเมินอภิमानที่พัฒนาขึ้น

6. คัดเลือกนักประเมินอภิमान จากผู้ที่มีความรู้ประสบการณ์ในการประเมินคุณภาพภายใน ระดับอุดมศึกษา หรือผ่านการอบรมหลักสูตรผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษา ของสำนักงานคณะกรรมการการอุดมศึกษาแล้ว จำนวน 28 คน รวม 2 รุ่น รุ่นแรกเป็นผู้ประเมินคุณภาพภายในของสถาบันอุดมศึกษา และผู้ที่เกี่ยวข้องกับการประเมินคุณภาพภายในของสถาบันอุดมศึกษา โดยทำการอบรมที่คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ระหว่างวันที่ 19-20 กันยายน 2552 จำนวน 12 คน ส่วนรุ่นที่ 2 เป็นผู้ประเมินคุณภาพภายในของมหาวิทยาลัยศรีปทุม และมหาวิทยาลัยอื่น โดยทำการฝึกอบรมที่ห้องบงกช อาคาร 10 ชั้น 6 มหาวิทยาลัยศรีปทุม ระหว่างวันที่ 4-5 มกราคม 2553 จำนวน 16 คน

7. นำนักประเมินอภิमानที่คัดเลือกตามคุณสมบัติที่กำหนด มาฝึกอบรมในหลักสูตร “การฝึกอบรมพัฒนานักประเมินอภิमान” ตามขั้นตอน วิธีการประเมินอภิमानที่เป็นตามมาตรฐานการประเมินอภิमानที่พัฒนาขึ้น

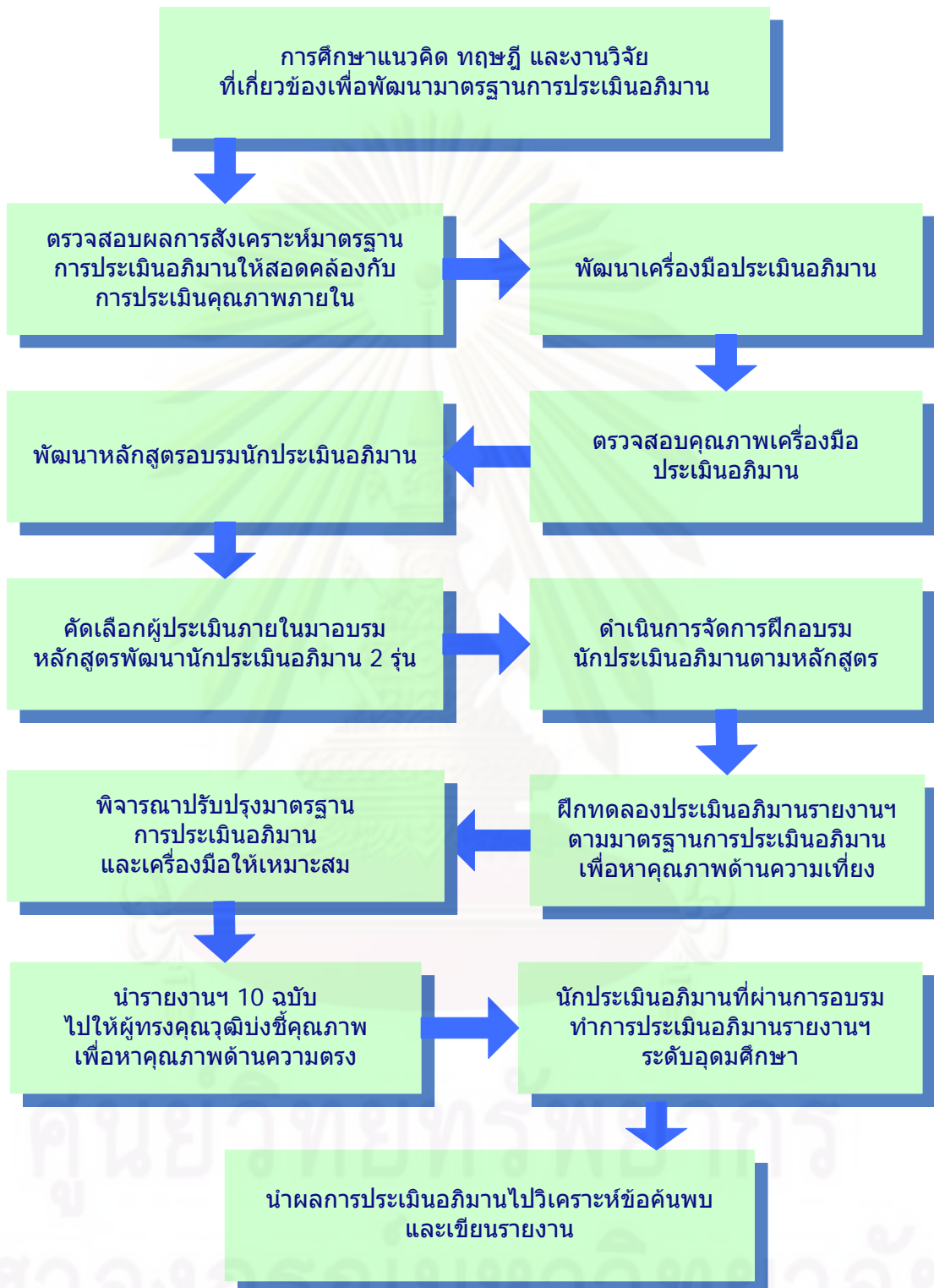
8. ฝึกประเมินอภิमानรายงานการประเมินคุณภาพของสถาบันอุดมศึกษา 5 ฉบับ ตามขั้นตอนการฝึกอบรมนักประเมินอภิमान โดยผู้เข้าอบรม ทำการประเมินรายงานผลการประเมินของสถาบันอุดมศึกษาจำนวน 2 ฉบับไขว้กัน เพื่อเป็นการตรวจสอบไขว้ผลการประเมิน และทำการวิเคราะห์หาคุณภาพด้านความเที่ยง (reliability) แบบสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient)

9. นำผลการตรวจสอบคุณภาพมาตรฐานการประเมินอภิमानและเครื่องมือประเมิน อภิमान ที่ผ่านการทดลองประเมิน มาพิจารณาปรับปรุงเครื่องมือในการประเมินอภิमानให้เหมาะสม มากยิ่งขึ้น

10. นำรายงานผลการประเมินคุณภาพภายใน จำนวน 10 ฉบับ ไปให้ผู้ทรงคุณวุฒิ ด้านการประกันคุณภาพจำนวน 10 ท่าน ประเมินคุณภาพรายงานฯ เพื่อบ่งชี้ระดับคุณภาพ เป็นการ พิจารณาคุณภาพด้านความตรงเชิงโครงสร้าง (construct validity) ด้วยวิธีเปรียบเทียบกับกลุ่มผู้ชัด (known group method) โดยผู้ทรงคุณวุฒิด้านการประกันคุณภาพ ระดับอุดมศึกษา เป็นผู้บ่งชี้คุณภาพ ของรายงานผลการประเมินคุณภาพภายใน โดยใช้เครื่องมือแบบประเมินคุณภาพรายงาน ที่พัฒนาขึ้น จาก นางลักษณะ วิรัชชัย และวรรณิ์ เจตจำนงนุช (2548) และผ่านการพิจารณาคุณภาพ ด้วยการ ตรวจสอบความสอดคล้องระหว่างอาจารย์ที่ปรึกษางานวิจัย และผู้วิจัย เพื่อพิจารณาหาคุณภาพด้าน ความเที่ยง (inter-rater reliability) เพื่อนำผลการประเมินจากผู้ทรงคุณวุฒิ มาเปรียบเทียบกับผลการ ประเมินอภิमानกับกลุ่มนักประเมินอภิमानที่พัฒนาขึ้น

11. นักประเมินอภิमानที่ผ่านการอบรมในหลักสูตร การฝึกอบรมพัฒนานักประเมิน อภิमान มีความพร้อมที่จะสามารถทำการประเมินอภิमानได้ ดำเนินการประเมินอภิमानรายงานผล การประเมินคุณภาพภายใน ของสถานศึกษา ระดับอุดมศึกษา จำนวน 45 ฉบับ ตามขั้นตอน กระบวนการที่ได้รับการฝึกอบรม เพื่อพิจารณาหาคุณภาพของรายงานผลการประเมินคุณภาพ ภายใน ของสถานศึกษา ตามมาตรฐานการประเมินอภิमानที่พัฒนาขึ้น

12. นำผลการประเมินอภิमानที่ได้จากการประเมินไปรวบรวม วิเคราะห์ข้อค้นพบทั้ง เชิงปริมาณ และเชิงคุณภาพ ดำเนินการสรุปผลข้อค้นพบและเขียนรายงาน ดังภาพที่ 3.1



ภาพที่ 3.1 ขั้นตอนการดำเนินการวิจัย

2. แหล่งข้อมูลสำหรับการวิจัย

การวิจัยในครั้งนี้มุ่งพัฒนามาตรฐานการประเมินอภิมาน เพื่อใช้ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา ระดับอุดมศึกษาตามมาตรฐานการประเมินอภิมานที่พัฒนาขึ้น และสังเคราะห์รายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา เพื่อให้ได้ข้อค้นพบที่เป็นประโยชน์ต่อการปรับปรุงระบบการประเมินคุณภาพ ระดับอุดมศึกษา และเพื่อให้ได้ข้อเสนอแนะสำหรับการจัดการประเมินคุณภาพภายในของสถาบันอุดมศึกษา จึงกำหนดแหล่งข้อมูลสำหรับการวิจัยออกเป็น 3 กลุ่ม ดังต่อไปนี้

1. ตัวอย่างรายงานผลการประเมินคุณภาพภายในของสถานศึกษา และรายงานการประเมินตนเอง ระดับอุดมศึกษา ปีการศึกษา 2550 จำนวน 50 แห่ง ที่มาจากสถาบันอุดมศึกษาที่สังกัด สำนักงานคณะกรรมการการอุดมศึกษาทั้งสิ้น 143 แห่ง ประกอบด้วยมหาวิทยาลัยของรัฐ 29 แห่ง มหาวิทยาลัยเอกชน 65 แห่ง มหาวิทยาลัยราชภัฏ 40 แห่ง และมหาวิทยาลัยเทคโนโลยีราชมงคล 9 แห่ง โดยกลุ่มตัวอย่างที่สุ่มเลือกมา มีเงื่อนไขดังต่อไปนี้

1.1) ต้องได้รับความยินยอม ในการนำข้อมูลไปใช้จากหน่วยงานต้นสังกัด คือ สำนักงานคณะกรรมการการอุดมศึกษา และสถาบัน

1.2) ต้องได้รับการประเมินคุณภาพภายใน ระดับสถาบัน และมีการเผยแพร่ โดยมีข้อมูลครบถ้วน สมบูรณ์ ทั้ง 2 ส่วน คือ รายงานการประเมินตนเอง (SAR) และรายงานผลการประเมินคุณภาพภายใน ประจำปีการศึกษา 2550 จึงจะสามารถนำมาประเมินอภิมานได้

ดังนั้นจึงทำการคัดเลือกตัวอย่างแบบเจาะจง (purposive sampling) ตามเงื่อนไขที่กำหนดในข้างต้น ทำการรวบรวมรายงานการประเมินตนเอง และรายงานผลการประเมินคุณภาพภายใน จากการ Download รายงานจาก website ของมหาวิทยาลัย หากไม่มีเผยแพร่ใน website จึงส่ง CD ทางไปรษณีย์เพื่อขอไฟล์รายงาน และนอกจากนี้ดำเนินการขอสำเนารายงานฯ จากสำนักงานคณะกรรมการการอุดมศึกษา ดังนั้น จึงได้รายงานฯ ครบทั้ง 50 แห่ง กลุ่มตัวอย่างเป็น มหาวิทยาลัยของรัฐ จำนวน 8 แห่ง มหาวิทยาลัยเอกชน จำนวน 19 แห่ง มหาวิทยาลัยราชภัฏ จำนวน 18 แห่ง และมหาวิทยาลัยเทคโนโลยีราชมงคล จำนวน 5 แห่ง ตามตารางที่ 3.1

ตารางที่ 3.1 จำนวนตัวอย่างรายงานผลการประเมินคุณภาพภายใน ระดับสถาบัน

สังกัดของสถาบันอุดมศึกษา	ประชากร	กลุ่มตัวอย่าง
มหาวิทยาลัยของรัฐ	29	8
มหาวิทยาลัยเอกชน	65	19
มหาวิทยาลัยราชภัฏ	40	18
มหาวิทยาลัยเทคโนโลยีราชมงคล	9	5
รวม	143	50

2. นักประเมินอภิमान (meta-evaluator) เป็นผู้ที่เข้ารับการฝึกอบรมหลักสูตรพัฒนานักประเมินอภิमानทั้ง 2 รุ่น จากรุ่นที่ 1 จำนวน 12 คน และรุ่นที่ 2 จำนวน 16 คน โดยผู้เข้ารับการอบรมต้องมีคุณสมบัติ เป็นผู้ประเมินคุณภาพภายใน หรือมีความรู้ ความสามารถเกี่ยวกับการประกันคุณภาพหรือการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา และเข้ารับการอบรมครบตามกำหนด 2 วัน เข้าร่วมกิจกรรมการฝึกอบรมตามหลักสูตร มีความรู้ ความเข้าใจในหลักการมาตรฐานการประเมินอภิमान และสามารถทำการประเมินอภิमानรายงานตามที่กำหนดได้ จากนั้นดำเนินการคัดเลือกนักประเมินอภิमानที่ผ่านการฝึกอบรมตามหลักสูตร จำนวน 15 คน ที่มีความพร้อมที่จะทำหน้าที่เป็นนักประเมินอภิमान โดยสามารถทำการประเมินอภิमानรายงานผลการประเมินคุณภาพภายในตามที่ได้รับมอบหมาย และเขียนรายงานการประเมินอภิमानตามที่กำหนดได้

3. ผู้ทรงคุณวุฒิด้านการประกันคุณภาพ จำนวน 10 ท่าน ดังรายชื่อในภาคผนวก ก ที่มีคุณสมบัติดังนี้ คือ มีประสบการณ์ด้านการประเมินคุณภาพการศึกษาอย่างน้อย 5 ปีขึ้นไป หรือมีวุฒิการศึกษาด้านการวัดผล ประเมินผล วิจัยการศึกษาหรือสาขาอื่นที่เกี่ยวข้อง หรือมีผลงานวิจัยทางด้านการประเมิน/การประกันคุณภาพการศึกษาอย่างน้อย 1 ชิ้น และเป็นผู้ประเมินสถานศึกษาระดับอุดมศึกษา เพื่อทำหน้าที่เป็นผู้บ่งชี้คุณภาพของรายงานผลการประเมินคุณภาพภายใน สำหรับการวิเคราะห์ความตรงเชิงโครงสร้าง (construct validity) ด้วยเทคนิคกลุ่มรู้ชัด (known group technique)

3. เครื่องมือและการพัฒนาเครื่องมือในการวิจัย

ขั้นตอนในการพัฒนาเครื่องมือในการวิจัย ประกอบด้วย 2 ส่วน คือ 3.1 เครื่องมือที่ใช้ในการวิจัย และ 3.2 การพัฒนาเครื่องมือในการวิจัย ดังมีรายละเอียดดังต่อไปนี้

3.1 เครื่องมือที่ใช้ในการวิจัย ประกอบด้วยเครื่องมือรวมทั้งหมด 4 ชนิด คือ 1. แบบประเมินอภิमान 2. คู่มือการประเมินอภิमान และเกณฑ์การให้คะแนน 3. หลักสูตรการฝึกอบรมนักประเมินอภิमान และ 4. แบบประเมินคุณภาพงานประเมินสำหรับผู้ทรงคุณวุฒิ ดังมีรายละเอียดดังต่อไปนี้

1. **แบบประเมินอภิमान** เป็นเครื่องมือสำหรับประเมินคุณภาพงานประเมินที่ดำเนินการสร้างตามกรอบมาตรฐานการประเมินอภิमानที่พัฒนาขึ้น 5 มาตรฐาน คือ 1) ความตรงของการประเมิน 2) อรรถประโยชน์ 3) จริยธรรมในการประเมิน 4) ความน่าเชื่อถือของนักประเมิน และ 5) ประสิทธิภาพต้นทุน เพื่อใช้ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน

ระดับอุดมศึกษา ปีการศึกษา 2550 เป็นชนิดกำหนดเกณฑ์การให้คะแนน (rubric scoring) 4 ระดับ ดังนี้ คือ

4	หมายถึง	ระบุการดำเนินการได้ครบถ้วนตามมาตรฐาน และมีหลักฐานปรากฏ
3	หมายถึง	มีรายละเอียดการดำเนินการจริง แต่ยังไม่ครอบคลุมมาตรฐาน
2	หมายถึง	เริ่มดำเนินการในบางส่วน หรือดำเนินการแล้ว แต่ไม่ระบุรายละเอียด
1	หมายถึง	ยังไม่มี การดำเนินการ หรือ ดำเนินการไม่ถูกต้อง

โดยมีเกณฑ์ในการแปลความหมายระดับคุณภาพ ดังนี้

3.26 ขึ้นไป	หมายถึง	งานประเมินมีคุณภาพดีมาก
2.51 – 3.25	หมายถึง	งานประเมินมีคุณภาพดี
1.76 – 2.50	หมายถึง	งานประเมินมีคุณภาพพอใช้
น้อยกว่า 1.76	หมายถึง	ควรปรับปรุงคุณภาพงานประเมิน

สำหรับเกณฑ์ในการแปลความหมายระดับคุณภาพ อ้างอิงมาจากเกณฑ์การประเมินของศูนย์ส่งเสริมและพัฒนาพลังแผ่นดินเชิงคุณธรรม (ศูนย์คุณธรรม) สำนักงานบริหารและพัฒนาองค์ความรู้ (องค์การมหาชน) (2551)

แบบประเมินอภิมานประกอบด้วย 3 ตอน ดังนี้ คือ

ตอนที่ 1 เป็นข้อมูลทั่วไปของสถาบันอุดมศึกษา เป็นชนิดเติมคำ และแบบตรวจสอบรายการ ประกอบด้วย ชื่อมหาวิทยาลัย ประเภทของสถาบันอุดมศึกษา และที่ตั้งสถาบันอุดมศึกษา

ตอนที่ 2 แบบประเมินอภิมาน เป็นมาตรฐานประมาณค่า 4 ระดับ ชนิดกำหนดเกณฑ์การให้คะแนน (rubric scoring) ที่ประกอบด้วย 5 มาตรฐาน และ 38 ตัวบ่งชี้

ดังตัวอย่าง ตัวบ่งชี้ Va 1.7 การตรวจสอบความถูกต้องของข้อมูล

เกณฑ์การประเมิน

1. ไม่ระบุการตรวจสอบความถูกต้องของข้อมูล
2. มีการระบุการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา แต่ไม่มีหลักฐานยืนยันแสดงความสอดคล้องกันของข้อมูล
3. มีการระบุการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา แต่หลักฐานยืนยันความสอดคล้องกันของข้อมูลต่างๆ ยังไม่มีรายละเอียด
4. มีการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา และมีหลักฐานยืนยันความสอดคล้องกันของข้อมูลต่างๆ มีรายละเอียดชัดเจน

ตอนที่ 3 ข้อเสนอแนะในการนำผลการประเมินไปใช้ประโยชน์ เป็นชนิดแบบสอปลายเปิด ดังรายละเอียดในภาคผนวก ข

2. คู่มือการประเมินอภิमान และเกณฑ์การให้คะแนน เป็นคำอธิบายแนวทางสำหรับการประเมินอภิमान มาตรฐานและเกณฑ์ที่ใช้ในการประเมินอภิमान โดยมีส่วนประกอบ 4 ส่วน ดังมีรายละเอียดในภาคผนวก ข ดังต่อไปนี้

ส่วนที่ 1 แนวคิดและความเป็นมาของหลักสูตรฝึกอบรมนักประเมินอภิमान

ส่วนที่ 2 คำอธิบายมาตรฐานการประเมินอภิमानในแต่ละมาตรฐาน และเกณฑ์ในการตัดสินการประเมิน

ส่วนที่ 3 บทบาทหน้าที่ของนักประเมินอภิमान และวิธีประเมินอภิमान

ส่วนที่ 4 ส่วนประกอบของรายงานการประเมินอภิमान

การจัดทำคู่มือการประเมินอภิमान เกณฑ์การให้คะแนน ผู้วิจัย ศึกษาจากคู่มือการประกันคุณภาพภายใน ระดับอุดมศึกษา และคู่มือการประเมินคุณภาพภายนอก ระดับอุดมศึกษา ดำเนินการจัดทำร่างคู่มือการประเมินอภิमान และนำไปให้ผู้เชี่ยวชาญทางด้านการประเมินพิจารณาความถูกต้อง เหมาะสมของคู่มือ ดำเนินการปรับปรุงคู่มือให้เหมาะสมก่อนนำไปใช้ รายละเอียดปรากฏในภาคผนวก ข

3. หลักสูตรการฝึกอบรมนักประเมินอภิमान เป็นหลักสูตรที่ใช้ในการฝึกอบรมนักประเมินอภิमान ให้มีความรู้ ทักษะที่สามารถทำการประเมินอภิमानได้ โดยแนวทางการพัฒนาหลักสูตรการฝึกอบรม ศึกษาจากการจัดทำหลักสูตร การฝึกอบรมผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษา ของสำนักงานคณะกรรมการการอุดมศึกษา หลักสูตรการฝึกอบรมผู้ประเมินอภิमान และสอบถามข้อมูลการจัดการฝึกอบรมจากหน่วยงานที่ดำเนินการจัดการฝึกอบรมผู้ประเมินอภิमान ได้แก่ ศูนย์เครือข่าย สมศ. จุฬาลงกรณ์มหาวิทยาลัย เพื่อจัดทำร่างหลักสูตรการฝึกอบรมนักประเมินอภิमान และตรวจสอบคุณภาพ โดยให้อาจารย์ที่ปรึกษา และผู้เชี่ยวชาญพิจารณาความเหมาะสมของหลักสูตรจัดฝึกอบรมนักประเมินอภิमान นอกจากนี้เข้าร่วมสังเกตการณ์ การจัดฝึกอบรมผู้ประเมินอภิमान ขึ้นพื้นฐานที่จัดขึ้นโดยศูนย์เครือข่าย สมศ. ในวันที่ 3 สิงหาคม 2552 ณ โรงแรมแมนดาริน จากนั้นนำหลักสูตรการฝึกอบรมมาพิจารณาปรับปรุงตามข้อเสนอแนะ เพื่อจัดทำหลักสูตรการฝึกอบรมที่เหมาะสมต่อไป ซึ่งประกอบด้วย

3.1 หลักการและเหตุผลในการฝึกอบรม

3.2 วัตถุประสงค์ของหลักสูตร 1) เพื่อสร้างความรู้ ความเข้าใจ และบทบาทหน้าที่ของนักประเมินอภิमान 2) เพื่อให้มีความเข้าใจในมาตรฐานและเกณฑ์ในการประเมินอภิमान 3)

เพื่อให้มีทักษะในการประเมินอภิमान สามารถตรวจสอบกระบวนการประเมินคุณภาพภายใน และการจัดทำรายงานผลการประเมินคุณภาพภายในได้อย่างถูกต้อง สอดคล้องกับมาตรฐานการประเมิน อภิमानและแนวทางการประเมินของสกอ. 4) เพื่อให้ผู้ประเมินสามารถนำความรู้ที่ได้รับไปพัฒนา ส่งเสริมการประเมินอภิमानภายในหน่วยงาน

3.3 ผู้รับผิดชอบหลักสูตรอบรม โดยผู้วิจัย ร่วมกับสำนักงานแผนและประกัน คุณภาพ มหาวิทยาลัยศรีปทุม

3.3 วิทยากร โดยผู้วิจัย ร่วมกับวิทยากรรับเชิญที่มีประสบการณ์เป็นผู้ประเมิน คุณภาพภายใน ระดับอุดมศึกษา ของสำนักงานคณะกรรมการการอุดมศึกษา

3.4 วัน เวลา สถานที่ รุ่นที่ 1 ระหว่างวันที่ 19-20 กันยายน 2552 เวลา 08.30- 16.30 น. ณ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย รุ่นที่ 2 ระหว่างวันที่ 4 - 5 มกราคม 2553 เวลา 08.30-16.30 น. ณ ห้องบงกช อาคาร 10 ชั้น 6 มหาวิทยาลัยศรีปทุม

3.5 วิธีการฝึกอบรม เน้นการฝึกปฏิบัติ โดยการทำกิจกรรมกลุ่ม ช่วยกันอภิปรายผล เพื่อหาข้อเสนอนะ จากการอบรมแต่ละกลุ่มนำเสนอผลงานและช่วยกันสรุปองค์ความรู้แต่ละหัวข้อ

3.6 ประโยชน์ที่คาดว่าจะได้รับ ผู้เข้ารับการอบรมมีความรู้ความเข้าใจใน มาตรฐานและเกณฑ์การประเมินอภิमान สามารถดูแล กำกับ ติดตาม กระบวนการประเมิน โดย สามารถดำเนินการประเมินอภิमानภายนอก และสามารถนำความรู้ ทักษะที่ได้รับไปพัฒนางาน ประเมินคุณภาพภายในและประเมินอภิमानของสถานศึกษาได้

3.7 คุณสมบัติของผู้เข้าอบรม ต้องเคยเข้าร่วมอบรมหลักสูตรผู้ประเมินคุณภาพ ภายใน ระดับอุดมศึกษา ของสกอ. หรือ หลักสูตรอื่นๆ ที่เกี่ยวข้องกับการประเมินคุณภาพ การศึกษา ระดับอุดมศึกษา หรือเป็นผู้มีความรู้ และประสบการณ์ในการประกันคุณภาพ หรือ ประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา

3.8 จำนวนผู้เข้าร่วมอบรม รุ่นละ 10-15 คน

3.9 เกณฑ์การในการผ่านการฝึกอบรม ต้องเข้าร่วมอบรมไม่น้อยกว่าร้อยละ 80 ของกิจกรรมระหว่างฝึกอบรม และต้องมีผลงานหลังฝึกอบรมตามเกณฑ์

4. แบบประเมินคุณภาพรายงานประเมิน สำหรับผู้ทรงคุณวุฒิ เป็นเครื่องมือ ที่ใช้ในการบ่งชี้คุณภาพของรายงานผลการประเมินคุณภาพภายใน จากการประเมินรายงานผลการ ประเมินคุณภาพภายใน โดยผู้ทรงคุณวุฒิ ประกอบด้วยข้อคำถาม 20 ข้อ ดังรายละเอียดใน ภาคผนวก ข โดยมีเกณฑ์การพิจารณาชนิดกำหนดเกณฑ์การให้คะแนน (rubric scoring) 5 ระดับ ดังต่อไปนี้

4	หมายถึง	ระบुरายละเอียดการดำเนินการได้ครบถ้วนตามมาตรฐาน และมีหลักฐานปรากฏ
3	หมายถึง	ดำเนินการได้ถูกต้อง ชัดเจน แต่ยังไม่ครอบคลุมตามมาตรฐาน
2	หมายถึง	ดำเนินการได้มีรายละเอียด บางส่วน
1	หมายถึง	เริ่มดำเนินการ แต่ไม่ถูกต้อง
0	หมายถึง	ยังไม่มีดำเนินการ

โดยมีเกณฑ์ในการแปลผลการประเมินคุณภาพงานประเมิน ดังนี้

คะแนนประเมินอยู่ระหว่าง 3.21 – 4.00	หมายถึง	คุณภาพดีมาก
คะแนนประเมินอยู่ระหว่าง 2.41 – 3.20	หมายถึง	คุณภาพดี
คะแนนประเมินอยู่ระหว่าง 1.61 – 2.40	หมายถึง	คุณภาพปานกลาง
คะแนนประเมินอยู่ระหว่าง 0.81 – 1.60	หมายถึง	คุณภาพค่อนข้างต่ำ
คะแนนประเมินต่ำกว่า 0.80	หมายถึง	คุณภาพต่ำ

3.2 การพัฒนาเครื่องมือในการวิจัย ผู้วิจัยขอเสนอขั้นตอนในการพัฒนาเครื่องมือในการวิจัย 2 ชนิด คือ 1. แบบประเมินอภิमान 2. แบบประเมินคุณภาพรายงานประเมิน โดยผู้ทรงคุณวุฒิ ดังมีรายละเอียดต่อไปนี้

1. การตรวจสอบคุณภาพแบบประเมินอภิमान แบ่งออกเป็น 3 ขั้นตอน คือ 1) การตรวจสอบคุณภาพด้านความตรงตามเนื้อหา(content validity) ด้วยการคำนวณหาค่าดัชนีความสอดคล้อง (indexes of Item - Objective Congruence : IOC) 2) การทดลองใช้และปรับปรุงแบบประเมินอภิमान และ 3) การตรวจสอบคุณภาพด้านความเที่ยง (reliability) ของแบบประเมินอภิमान โดยวิธีสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach' s alpha) ดังมีรายละเอียดต่อไปนี้

1) การตรวจสอบคุณภาพด้านความตรงตามเนื้อหา(content validity) ด้วยการคำนวณหาค่าดัชนีความสอดคล้อง (indexes of Item - Objective Congruence : IOC) ของแบบประเมินอภิमान จากการพิจารณาของผู้เชี่ยวชาญทางด้านการประกันคุณภาพ ด้านการวัดผลและด้านการประเมินผล จำนวน 5 ท่าน โดยตรวจสอบความสอดคล้องของเนื้อหา ข้อความ และภาษาที่ใช้ในแบบประเมินว่า วัดได้ตรงคุณลักษณะที่ต้องการหรือไม่ หากแน่ใจว่าวัดได้สอดคล้อง ให้คะแนน +1 หากแน่ใจว่าวัดได้ไม่สอดคล้องให้คะแนน -1 หากไม่แน่ใจให้คะแนน 0 แล้วนำผลคะแนนที่ได้จากการพิจารณาของผู้เชี่ยวชาญ ไปคัดเลือกข้อคำถามที่มีค่าดัชนีความสอดคล้องตั้งแต่ .50 ขึ้นไป (Rovinelli & Hambleton (1977; Turner & Carlson, 2003 อ้างถึงใน Rovinelli & Hambleton, 1977)

มีผลการวิเคราะห์ดังตารางที่ 3.2 ซึ่งพบว่า มีข้อที่ค่า IOC เท่ากับ 0.4 ต่ำกว่าเกณฑ์ 0.5 จำนวน 5 ข้อ จึงนำไปพิจารณาปรับปรุงข้อความ และเกณฑ์ในการประเมินให้เหมาะสมในการวัดมากขึ้น และนำไปทดลองใช้อีกครั้งหนึ่ง

ตารางที่ 3.2 ผลการวิเคราะห์ความตรงตามเนื้อหา (IOC) ของแบบประเมินอภิमानจากผู้เชี่ยวชาญ

มาตรฐาน ที่	มาตรฐานการประเมิน	ค่า IOC	หมายเหตุ
1.	ด้านความตรงของการประเมิน (Validity)	0.4-1.0	Va 4.1 (0.4) ปรับปรุง
2.	ด้านอรรถประโยชน์ (Utility)	0.4-1.0	Ut 2.1 (0.4) ปรับปรุง
3.	ด้านจริยธรรมในการประเมิน (Ethicality)	0.4-0.8	Et 3.1, 7.1, 9.1 ปรับปรุง
4.	ด้านความน่าเชื่อถือของนักประเมิน (Credibility)	0.8	
5.	ด้านประสิทธิผลต้นทุน (Cost-effectiveness)	0.8	
ค่า IOC มีค่าอยู่ระหว่าง		0.4-1.0	ปรับปรุง 5 ตัวบ่งชี้

2) การทดลองใช้และปรับปรุงแบบประเมินอภิमान โดยดำเนินการปรับปรุงแบบประเมินอภิमानรวม 3 ครั้ง ดังต่อไปนี้

ครั้งที่ 1 ดำเนินการปรับปรุงแบบประเมินอภิमानตามข้อเสนอแนะของผู้เชี่ยวชาญในการพิจารณาเครื่องมือ ปรับแก้ไขข้อตัวบ่งชี้ให้สื่อความหมายมากขึ้น ปรับปรุงเกณฑ์การประเมินให้เหมาะสมมากขึ้น

ครั้งที่ 2 ดำเนินการปรับปรุงหลังจากนำแบบประเมินอภิमानไปทดลองใช้กับกลุ่มผู้เข้าร่วมอบรมนักประเมินอภิमान รุ่น 1 ระหว่างวันที่ 19-20 กันยายน 2552 ณ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปรับปรุงประเด็นตามตัวบ่งชี้ ปรับปรุงเกณฑ์ให้ชัดเจนมากขึ้น เพิ่มนิยามคำศัพท์ในเกณฑ์การประเมิน

ครั้งที่ 3 ดำเนินการปรับปรุงหลังจากนำเครื่องมือไปใช้ในการอบรมนักประเมินอภิमान รุ่นที่ 2 ระหว่างวันที่ 4-5 มกราคม 2553 ณ มหาวิทยาลัยศรีปทุม ในประเด็นดังนี้ คือ ปรับปรุงเกณฑ์ในการประเมิน และนิยามคำศัพท์ให้เหมาะสม ชัดเจน

สรุปการปรับปรุงตัวบ่งชี้ และเกณฑ์ในการประเมิน ให้ชัดเจน เหมาะสม และวัดได้ เป็นรูปธรรมมากยิ่งขึ้น ดังรายละเอียดในตารางที่ 3.3

ตารางที่ 3.3 ข้อเสนอแนะในการปรับปรุงแบบประเมินอภิमान

ตัวบ่งชี้	ข้อเสนอแนะเพื่อปรับปรุง		
	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
Va1.1 การระบุบริบทของสถาบัน	แก้ไขเกณฑ์ให้ชัดเจนขึ้น	ไม่ชัดเจนในบริบทของสถาบัน	
Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมิน	แก้ไขข้อบ่งชี้เป็น "การบรรยายถึงวัตถุประสงค์ในการประเมินอย่างโปร่งใส"	วัตถุประสงค์การประเมิน โปร่งใส ตรวจสอบยาก	
Va1.4 การออกแบบการประเมิน	แก้ไขข้อบ่งชี้เป็น "การออกแบบการประเมินที่สอดคล้องกับสภาพ และบริบทของสถาบัน" แก้ไขเกณฑ์		ปรับการเขียนเกณฑ์ให้ง่ายต่อความเข้าใจ
Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสาร	แก้ไขข้อบ่งชี้เป็น "การวิเคราะห์ความแตกต่างของเอกสารการประเมิน" และแก้ไขเกณฑ์		
Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	แก้ไขเกณฑ์ ที่ 1, 2 จำนวนแหล่งข้อมูล		
Va1.7 การตรวจสอบความถูกต้องของข้อมูล	แก้ไขเกณฑ์ ให้เหมาะสม	ร้อยละความถูกต้องวัดยาก	
Va1.8 การตรวจสอบคุณภาพของสารสนเทศ	แก้ไขเกณฑ์ให้เหมาะสม	คุณภาพเครื่องมือคืออะไร	เครื่องมือมีคุณภาพ ไม่ได้ระบุการตรวจสอบ
Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูล	แก้ไขเกณฑ์วัดจากจำนวนตัวบ่งชี้ที่วิเคราะห์ถูกต้อง		
Va1.10 การแปลผลและสรุปผลการประเมินถูกต้อง	แก้ไขเกณฑ์	ยังไม่ชัดเจนการสำรวจข้อบกพร่อง	
Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ	แก้ไขเกณฑ์ ให้เหมาะสม		
Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง	แก้ไขเกณฑ์ ให้เหมาะสม		
Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียในการประเมิน	แก้ไขเกณฑ์ จำนวนผู้มีส่วนได้ส่วนเสีย	ผู้มีส่วนได้ส่วนเสียใหม่ คือใครได้บ้าง	ควรนิยามว่าผู้มีส่วนได้ส่วนเสียมีส่วนร่วมแค่ไหน
Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล	แก้ไขเกณฑ์ ให้เหมาะสม กับตัวบ่งชี้		
Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้	สลับเกณฑ์ 3 กับ 4	
Ut 2.4 การตรวจสอบการตัดสินผลการประเมิน	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้		
Ut 2.5 ผลการประเมินเป็นประโยชน์ ในการพัฒนาการดำเนินงาน	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้	การนำผลไปใช้ เป็นของผู้ถูกประเมินมากกว่า เกณฑ์ ไม่สอดคล้องกับตัวบ่งชี้	ควรปรับปรุงเกณฑ์

ตารางที่ 3.3 (ต่อ)

ตัวบ่งชี้	ข้อเสนอแนะเพื่อปรับปรุง		
	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
Ut 2.6 การเขียนรายงาน สอดคล้องกับรูปแบบที่กำหนด	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้	ถูกต้องตามรูปแบบสกอ. ขอบเขตแค่นั้น	
Ut 2.7 รายงานผลการประเมิน ครบถ้วน และข้อเสนอแนะเพื่อ การพัฒนา	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้		
Ut 2.8 รายงานผลการประเมิน เขียนด้วยภาษาเชิงวิชาการ	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้		
Ut 2.9 การจัดทำรายงานผล การประเมินเสร็จทันเวลา	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้		ควรเปลี่ยนเป็นตรวจประเมิน ภายในกำหนดแทน
Ut 2.10 การเผยแพร่ผลการ ประเมินสู่ผู้ใช้ผลการประเมินทุก กลุ่ม	แก้ไขเกณฑ์ให้เหมาะสม กับตัวบ่งชี้	เป็นหน้าที่ของผู้ประเมินหรือ สถาบัน	ควรปรับปรุงเกณฑ์
Et 3.1 การสื่อสารกับผู้มีส่วน ร่วมในการประเมิน	แก้ไขข้อบ่งชี้เป็น "การสื่อสารกับผู้มี ส่วนร่วมในการประเมินตลอดช่วงเวลา ประเมิน" แก้ไขเกณฑ์	การสื่อสารตลอดเวลาดียก	
Et 3.2 ผลการประเมินเป็นที่ ยอมรับจากผู้ที่เกี่ยวข้อง	แก้ไขเกณฑ์ที่ 1 เป็นผลการประเมิน ไม่ได้รับการยอมรับ		บุคคลหลายกลุ่ม หมายความว่า อย่างไร และกรณีมีการทักท้วง แต่ไม่ได้ระบุไว้ในรายงาน จะ ไม่สามารถประเมินได้ ปราศจากอคติดูอย่างไร
Et 3.3 การปรับปรุงคุณภาพการ ประเมินอย่างต่อเนื่อง	ปรับปรุงเกณฑ์ให้สอดคล้องกับตัวบ่งชี้	เกณฑ์ไม่สอดคล้องกับ มาตรฐาน ตัวบ่งชี้	ปรับเกณฑ์ให้เหมาะสมกับ บริบทการเขียนรายงานการ ประเมิน
Et 3.5 การเปิดเผยผลการ ประเมินและข้อจำกัด	ปรับปรุงเกณฑ์ให้สามารถวัดได้จาก เอกสารรายงาน		ไม่มีกำหนดไว้ในรูปแบบ รายงานการประเมิน
Et 3.6 การกำหนดวิธีการ ประเมินโดยไม่ก่อให้เกิดความ เสียหายกับผู้ให้ข้อมูล			เกณฑ์ข้อ 3 และ 4 ไม่ชัดเจน
Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย	ปรับแก้ไขเกณฑ์ให้สอดคล้องกับตัวบ่งชี้	ผู้เกี่ยวข้อง น่าจะเป็นทีมผู้ ประเมินเท่านั้น	ควรใช้คำว่า "การสัมภาษณ์ให้ ได้ข้อมูล" มากกว่าใช้ "การศึกษามุมหลัง" ปรับปรุง เกณฑ์
Cr4.1 ผู้ประเมินมี ความสามารถ ในการประเมิน		เกณฑ์ในการประเมินควรมี รายละเอียดที่ชัดเจน เหมือนกัน และให้ข้อมูลที่ ลดหลั่นลงมาในการให้ คะแนน	ควรปรับปรุงเกณฑ์
Cr 4.2 ผู้ประเมินมีทักษะในการ สื่อสารผลการประเมิน	ปรับปรุงเกณฑ์ให้เป็นรูปธรรม		

ตารางที่ 3.3 (ต่อ)

ตัวบ่งชี้	ข้อเสนอแนะเพื่อปรับปรุง		
	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมิน			เกณฑ์ในการพิจารณาควรปรับลดหลั่น จากปรับปรุง พอใช้ ดี ดีมาก
Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน	ปรับปรุงเกณฑ์ในการประเมินให้สอดคล้องกับการประเมิน		เกณฑ์ที่ระบุพิจารณายาก เข้าใจยาก ทำให้ประเมินไม่ได้ และไม่ได้ระบุไว้ในรูปแบบรายงานการประเมิน
Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า	ปรับปรุงเกณฑ์ให้เป็นรูปธรรม		การพิจารณาหลักฐานว่า มี/ไม่มีทำได้อาก
Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน			ไม่ได้ระบุไว้ในรูปแบบรายงานการประเมิน ปรับปรุงเกณฑ์

3) การตรวจสอบคุณภาพด้านความเที่ยงโดยนำแบบประเมินอภิมาน ไปการทดลองเก็บรวบรวมข้อมูลจากรายงานผลการประเมินคุณภาพภายใน โดยกลุ่มผู้เข้าร่วมฝึกอบรมนักประเมินอภิมาน จำนวน 2 รุ่น ทำการประเมินอภิมานรายงานผลการประเมินคุณภาพภายในฉบับเดียวกัน รวม 5 ฉบับ เพื่อนำมาหาคุณภาพด้านความเที่ยง (reliability) โดยวิธีสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach' s alpha) ของแบบประเมินอภิมาน มีค่าความเที่ยงทั้งฉบับเท่ากับ 0.788 ดังตารางที่ 3.4 และนำเครื่องมือมาปรับปรุงแก้ไขให้เหมาะสม

ตารางที่ 3.4 ผลการวิเคราะห์ความเที่ยงของแบบประเมินอภิมานสัมประสิทธิ์แอลฟา

มาตรฐานที่	มาตรฐานการประเมิน	จำนวนตัวบ่งชี้	ค่าความเที่ยง (reliability)	Standard Error of Measurement (SEM)
1.	ความตรงของการประเมิน (Validity)	13	0.511	0.265
2.	อรรถประโยชน์ (Utility)	10	0.739	0.303
3.	จริยธรรมในการประเมิน (Ethicality)	9	0.852	0.312
4.	ความน่าเชื่อถือของนักประเมิน (Credibility)	4	0.504	0.260
5.	ประสิทธิผลต้นทุน (Cost-effectiveness)	2	0.911	0.211
	ค่าความเที่ยงทั้งฉบับ	38	0.788	0.283

2. แบบประเมินคุณภาพรายงานประเมิน โดยผู้ทรงคุณวุฒิ ปรับปรุงพัฒนามาจากแบบประเมินคุณภาพรายงานประเมิน จากงานวิจัยของนางลักษณ์ วิรัชชัย และวรรณี เจตจำนงนุช (2548) ตรวจสอบคุณภาพโดยผู้วิจัย และอาจารย์ที่ปรึกษาวิทยานิพนธ์ ทดลองประเมินรายงานผลการประเมินคุณภาพภายใน ฉบับเดียวกันด้วยแบบประเมินที่ปรับปรุงขึ้น เพื่อวิเคราะห์การหาความสอดคล้องของแบบประเมินคุณภาพรายงานประเมินของผู้ทรงคุณวุฒิ ระหว่างอาจารย์ที่ปรึกษาวิทยานิพนธ์ และผู้วิจัย (inter-rater reliability) พบว่ามีความสอดคล้อง 77.78% หรือ 0.78 ซึ่งมีค่าสูงกว่า 0.75 และวิเคราะห์ตามวิธีของ Kappa Agreement ด้วยสูตรของ Cohen (1960) มีค่า .610 ที่ระดับนัยสำคัญ .01 แสดงว่ามีความสอดคล้องกัน

4. การเก็บรวบรวมข้อมูล

การเก็บรวบรวมข้อมูล เป็นออกเป็น 4 ระยะ ดังต่อไปนี้

ระยะแรก ขั้นพัฒนามาตรฐาน และเกณฑ์ในการประเมินอภิมาน แบ่งการดำเนินงานออกเป็น 5 ขั้นตอน ดังต่อไปนี้

1. การรวบรวมข้อมูลจากเอกสารและงานวิจัยที่เกี่ยวข้องกับการประเมินอภิมาน มาตรฐานการประเมิน พระราชบัญญัติการศึกษาแห่งชาติ มาตรฐานการศึกษา มาตรฐานการประเมิน 1) Program Evaluation Standards 2) The African Evaluation guidelines 2002 3) SEVAL's Evaluation Standards (2004) 4) The DeGEval's Standards (2003) 5) Guiding Principles for Evaluators ของ American Evaluation Association [AEA] (2003) 6) Essential Skills Series in Evaluation ของ Canadian Evaluation Society (1999) 7) Essential Competencies for Program Evaluators (Ghere & et al., 2006) และเอกสารอื่นๆที่เกี่ยวข้อง เพื่อทำการสังเคราะห์และพัฒนามาตรฐานการประเมินอภิมาน ใน 5 มาตรฐาน คือ มาตรฐานความตรงของการประเมิน (validity) มาตรฐานอรรถประโยชน์ (utility) มาตรฐานจริยธรรมในการประเมิน (ethicality) มาตรฐานความน่าเชื่อถือของนักประเมิน (credibility) และมาตรฐานประสิทธิผลต้นทุน (cost-effectiveness)

2. ดำเนินการนิยามเชิงปฏิบัติการ รายละเอียดในแต่ละมาตรฐานของการประเมิน และมาตรฐานการประเมิน โดยให้สอดคล้องกับสภาพและบริบทการประเมิน ในระดับอุดมศึกษาของไทย

3. พัฒนาตัวบ่งชี้ และเกณฑ์ในการประเมินให้สอดคล้องกับการประเมินในแต่ละมาตรฐาน เพื่อให้ให้นักประเมินอภิมานสามารถตัดสินผลการประเมินได้ตรงกัน

4. นำตัวบ่งชี้ และเกณฑ์ในการประเมินมาพัฒนาเป็นแบบประเมินอภิมานตามกรอบมาตรฐานการประเมินอภิมาน เพื่อใช้ในการประเมินคุณภาพรายงานการประเมินคุณภาพ

ภายใน ระดับอุดมศึกษา และนำไปให้ผู้เชี่ยวชาญทางด้านการวัดผลและประเมินผลและด้านการประกันคุณภาพ จำนวน 5 ท่าน พิจารณาหาคุณภาพด้านความตรงตามเนื้อหา (content validity) ของเครื่องมือ

5. นำผลจากการพิจารณาของผู้เชี่ยวชาญและข้อเสนอแนะ มาปรับปรุงแก้ไข แบบประเมินอภิमानให้เหมาะสมมากขึ้น

ระยะสอง **ขั้นการพัฒนาเครื่องมือการประเมินอภิमानและหลักสูตรการฝึกอบรมนักประเมินอภิमान**

แบ่งการดำเนินการออกเป็น 2 ส่วน คือ การพัฒนาเครื่องมือการประเมินอภิमान และหลักสูตรการฝึกอบรมนักประเมินอภิमान มีรายละเอียด ดังต่อไปนี้

ส่วนที่ 1 การพัฒนาเครื่องมือการประเมินอภิमान ประกอบด้วย 1) ดำเนินการศึกษาเอกสารที่เกี่ยวข้อง กับการพัฒนาเครื่องมือการประเมินคุณภาพภายใน ระดับอุดมศึกษา เครื่องมือการประเมินคุณภาพภายนอก ของ สมศ. รวมทั้งการประเมินอภิमान เพื่อกำหนดกรอบแนวในการวางขอบเขตเนื้อหาในเครื่องมือการประเมินอภิमान 2) ร่างเครื่องมือการประเมินอภิमानตามกรอบแนวทางในการจัดทำเครื่องมือการประเมินอภิमान 3) นำร่างเครื่องมือการประเมินอภิमानไปให้ผู้เชี่ยวชาญที่เกี่ยวข้อง พิจารณาความเหมาะสมและความเพียงพอของเนื้อหาในเครื่องมือ 4) พิจารณาปรับปรุงเครื่องมือการประเมินอภิमानให้เหมาะสมกับการประเมินอภิमान

ส่วนที่ 2 การพัฒนาหลักสูตรการฝึกอบรมนักประเมินอภิमान ประกอบด้วย 1) ดำเนินการศึกษาเอกสารที่เกี่ยวข้อง และหลักสูตรในการฝึกอบรมผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษา หลักสูตรฝึกอบรมผู้ประเมินอภิमान ขั้นพื้นฐาน ศึกษาจากหน่วยงานที่ทำหน้าที่ฝึกอบรมผู้ประเมินอภิमान ได้แก่ เครือข่าย สมศ. จุฬาลงกรณ์มหาวิทยาลัย เพื่อเป็นแนวทางในการจัดทำร่างหลักสูตรการฝึกอบรมนักประเมินอภิमान 2) จัดทำร่างหลักสูตรการฝึกอบรมนักประเมินอภิमानตามกรอบแนวทางที่ได้ศึกษาค้นคว้า โดยเข้าร่วมสังเกตการณ์การจัดการอบรมฝึกอบรมผู้ประเมินอภิमानขั้นพื้นฐาน ที่จัดโดย ศูนย์เครือข่าย สมศ. จุฬาลงกรณ์มหาวิทยาลัย ในวันที่ 3 สิงหาคม 2552 โดยมี ศาสตราจารย์ ดร. สุวิมล ว่องวาณิช เป็นวิทยากรอบรมการประเมินอภิमान เพื่อศึกษากระบวนการ ขั้นตอนการฝึกอบรมผู้ประเมินอภิमानเพื่อนำมาพิจารณาร่างหลักสูตรอบรมนักประเมินอภิमानต่อไป 3) นำร่างหลักสูตรการฝึกอบรมนักประเมินอภิमानไปให้อาจารย์ที่ปรึกษา และอาจารย์ที่ปรึกษาร่วม รวมถึงนำไปปรึกษาผู้มีความรู้ประสบการณ์ในการจัดการอบรมผู้ประเมินอภิमान พิจารณาความเหมาะสมของเนื้อหาและกิจกรรมในหลักสูตรการฝึกอบรม เพื่อ

เสนอแนะแนวทางในการพัฒนาปรับปรุงหลักสูตร 4) พิจารณาปรับปรุงหลักสูตรการฝึกอบรบตามข้อเสนอแนะให้เหมาะสมต่อไป

ระยะที่สาม ขั้นตอนการดำเนินการจัดฝึกอบรมนักประเมินอภิมาน

การดำเนินการจัดฝึกอบรมนักประเมินอภิมานตามมาตรฐานการประเมินอภิมาน โดยมีขั้นตอนการดำเนินงานดังต่อไปนี้ 1) จัดเตรียมเอกสารประกอบการฝึกอบรบประเมินอภิมาน ได้แก่ รายงานการประเมินตนเอง และรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา 5 แห่ง ที่ประกอบด้วย มหาวิทยาลัยของรัฐ มหาวิทยาลัยเอกชน 2 แห่ง มหาวิทยาลัยราชภัฏ และมหาวิทยาลัยเทคโนโลยีราชมงคล 2) ทำหนังสือเชิญเข้าร่วมฝึกอบรมนักประเมินอภิมาน โดยคัดเลือก ผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษาและผู้ที่มีความรู้ประสบการณ์ในการประกันคุณภาพ ระดับอุดมศึกษา ที่มีคุณสมบัติตามเกณฑ์ จำนวนรุ่นละ 10-15 คน 3) ดำเนินการจัดเตรียมประสานงานวิทยากรรับเชิญในการอบรม เอกสารประกอบการอบรม อาหาร สถานที่ รวมทั้งสิ่งอำนวยความสะดวกสำหรับจัดฝึกอบรบที่เหมาะสมทั้ง 2 รุ่น 4) ดำเนินการจัดฝึกอบรบตามหลักสูตรและวัน เวลาที่กำหนดไว้ 5) ให้นักประเมินอภิมานที่ฝึกอบรบทดลองประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน 5 ฉบับ ของสถานศึกษาที่คัดเลือกไว้ และให้ผู้เข้ารับการอบรมแต่ละคน ดำเนินการประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน 2 ฉบับ โดยทำการประเมินไขว้กัน เพื่อมาเปรียบเทียบหาคุณภาพด้านความเที่ยง (reliability) แบบสัมประสิทธิ์การสุร้อ้างอิง ของเครื่องมือ 6) สุ่มเลือกรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา จำนวน 10 ฉบับ เพื่อให้ผู้ทรงคุณวุฒิ จำนวน 10 ท่าน (รายนามดังภาคผนวก ก) ที่มีคุณสมบัติดังนี้ คือ มีประสบการณ์ด้านการประเมินคุณภาพอย่างน้อย 5 ปีขึ้นไป หรือ มีวุฒิการศึกษาด้านการวัดผล ประเมินผล วิจัยการศึกษาและสาขาอื่นๆที่เกี่ยวข้อง หรือ มีผลงานวิจัยทางด้านการประเมิน/การประกันคุณภาพอย่างน้อย 1 ชิ้น และเป็นผู้ประเมินสถานศึกษา ในระดับอุดมศึกษา ดำเนินการประเมินคุณภาพรายงานฯ ตามที่สุ่มเลือกมา เพื่อบ่งชี้คุณภาพของรายงานผลการประเมินคุณภาพภายใน 7) นำผลการพิจารณาของผู้ทรงคุณวุฒิที่ได้มาหาคุณภาพของมาตรฐานการประเมินอภิมานด้านความตรงเชิงโครงสร้าง (construct validity) ด้วยวิธีเปรียบเทียบกับกลุ่มรู้ชัด (known group method) และเพื่อนำไปเปรียบเทียบกับผลการประเมินอภิมานของนักประเมินอภิมานที่ประเมินคุณภาพรายงานฯ ฉบับเดียวกับผู้ทรงคุณวุฒิ ด้วยแบบประเมินอภิมาน และ 8) พิจารณาคัดเลือกนักประเมินอภิมานที่ผ่านเกณฑ์การฝึกอบรบ เพื่อทำหน้าที่เป็นนักประเมินอภิมานต่อไป

ระยะที่สี่ ขั้นตอนการประเมินอภิमान

การประเมินอภิमानมีขั้นตอนการดำเนินงานดังต่อไปนี้ คือ

1. ทำหนังสือติดต่อประสานงานกับสถาบันอุดมศึกษา และสำนักงานคณะกรรมการการอุดมศึกษา เพื่อขอข้อมูลรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา และรายงานการประเมินตนเองที่เป็นไปตามข้อกำหนด จำนวน 45 แห่ง
2. ดำเนินการประสานงานกับนักประเมินอภิमानที่ผ่านการฝึกอบรม เพื่อกำหนดภาระและแบ่งหน้าที่ในการประเมินอภิमानสถาบันอุดมศึกษา โดยให้นักประเมินอภิमानที่ผ่านการอบรมแล้ว ประเมินรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา คนละ 2-5 เล่ม โดยทำการประเมินรายงานผลการประเมินคุณภาพภายใน ต่างฉบับกัน จนครบ 45 แห่ง
3. นักประเมินอภิमानไปดำเนินประเมินอภิमानตามขั้นตอน การประเมินอภิमानที่ได้รับการฝึกอบรม และตามภาระหน้าที่ ที่ได้รับมอบหมายให้ประเมิน และส่งข้อมูลรายงานการประเมินอภิमानกลับมายังผู้วิจัย ตามที่ตกลงกัน
4. ผู้วิจัยรวบรวมข้อมูล จากรายงานการประเมินอภิमान และนำมาบันทึกข้อมูล และวิเคราะห์ข้อมูล เชิงปริมาณ และเนื้อหาต่อไป

5. การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลแบ่งออกเป็น 3 ส่วน ดังนี้

1. วิเคราะห์หาค่าสถิติพื้นฐาน สำหรับวิเคราะห์ผลการประเมินอภิमानรายงานการประเมินคุณภาพภายใน ระดับอุดมศึกษา ด้วยการหาค่าร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด
2. วิเคราะห์หาคุณภาพด้านความตรงของมาตรฐานการประเมินอภิमानด้วยการวิเคราะห์ความตรงเชิงเนื้อหา (content validity) จากการพิจารณาของผู้เชี่ยวชาญด้านการประกันคุณภาพ และวิเคราะห์หาความตรงเชิงโครงสร้าง (construct validity) ด้วยการวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) จากข้อมูลเชิงประจักษ์ ด้วยโปรแกรม LISREL 8.80 และนอกจากนี้ พิจารณาคุณภาพมาตรฐานการประเมินอภิमानด้วยการนำผลการประเมินอภิमानที่ได้จากการประเมินโดยนักประเมินอภิमानมาเปรียบเทียบกับคุณภาพของรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา ที่บ่งชี้คุณภาพโดยผู้เชี่ยวชาญด้านการประกันคุณภาพ ระดับอุดมศึกษา จำนวน 10 ท่าน และทำการวิเคราะห์ ด้วยสถิติทดสอบที (t-test)
3. วิเคราะห์หาคุณภาพด้านความเที่ยง (reliability) แบบสัมประสิทธิ์การสรูปอ้างอิง (generalizability coefficient) โดยนำผลการประเมินอภิमानจากนักประเมินอภิमानที่

ฝึกอบรมประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษาที่คัดเลือกไว้ และให้ผู้เข้ารับการอบรมแต่ละคน ดำเนินการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน โดยทำการประเมินไขว้กัน แบบ I X (R:P) Design เพื่อประมาณค่าความเที่ยง กรณีที่มีนักประเมินอภิमानตั้งแต่ 1, 2, และ 3 คน และประมาณค่ารายงานผลการประเมินคุณภาพภายใน จำนวน 3, 5, 7, และ 9 ฉบับ

สถิติที่ใช้ในการวิเคราะห์ข้อมูล

1. สถิติที่ใช้ในการวิเคราะห์หาค่าความตรงเชิงเนื้อหาของเครื่องมือประเมินอภิमान ด้วยค่าดัชนีความสอดคล้อง (indexes of Item - Objective Congruence: IOC) แล้วคัดเลือกข้อคำถามที่มีค่าดัชนีความสอดคล้องตั้งแต่ .50 ขึ้นไป Rovinelli & Hambleton (1977; Turner & Carlson, 2003 อ้างถึงใน Rovinelli & Hambleton, 1977)

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

เมื่อ IOC คือ ค่าดัชนีความสอดคล้อง

$\sum R$ คือ ผลรวมของคะแนนความคิดเห็นของผู้เชี่ยวชาญ

N คือ จำนวนผู้เชี่ยวชาญ

2. สถิติที่ใช้ในการคำนวณหาค่าความสอดคล้องระหว่างนักประเมินอภิमान พิจารณาค่าความสอดคล้องระหว่างนักประเมิน ตั้งแต่ .75 ขึ้นไป

$$\text{Agreement _ rate} = \frac{\text{จำนวนข้อที่ผู้ประเมินสองคนประเมินตรงกัน (AR)}}{\text{จำนวนข้อที่ประเมินทั้งหมด}}$$

และในกรณีข้อมูลเป็นระดับนามบัญญัติ และเรียงลำดับใช้สูตรการหาความสอดคล้องระหว่างผู้ประเมินด้วยสูตรของ Cohen (1960; Mcintire & Miller, 2007 อ้างถึงใน Cohen, 1960)

$$K = \frac{p_o - p_c}{1 - p_c}$$

เมื่อ K คือ สัดส่วนของความสอดคล้อง

p_o คือ สัดส่วนของคะแนนที่สังเกตได้

p_c คือ สัดส่วนของคะแนนที่คาดหวัง

3. สถิติที่ใช้คำนวณหาค่าความเที่ยงของแบบประเมินอภิมาน โดยใช้สัมประสิทธิ์ความสอดคล้องภายใน สูตรสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach' s alpha) (ศิริชัย กาญจนวาสี, 2548)

$$\sigma = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum s_i^2}{s_t^2} \right)$$

เมื่อ	α	คือ สัมประสิทธิ์ความเที่ยงแอลฟาของครอนบาค
	k	คือ จำนวนข้อความในแบบประเมินอภิมาน
	s_i^2	คือ ความแปรปรวนคะแนนประเมินรายข้อ
	s_t^2	คือ ความแปรปรวนคะแนนประเมินทั้งฉบับ

4. สถิติเปรียบเทียบความแตกต่างผลการประเมินอภิมานรายงาน เพื่อจำแนกคุณภาพรายงาน จำนวน 10 ฉบับ โดยใช้ t-test independent เพื่อหาคุณภาพด้านความตรงเชิงโครงสร้าง (construct validity) (ศิริชัย กาญจนวาสี, 2547)

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{S \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}, df = n_1 + n_2 - 2$$

$$\text{เมื่อ } S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

t	คือ สถิติที่ใช้ในการเปรียบเทียบค่าเฉลี่ยของผลการประเมินอภิมาน
\bar{X}_1	คือ ค่าเฉลี่ยของผลการประเมินอภิมานรายงานฯ กลุ่มคุณภาพสูง
\bar{X}_2	คือ ค่าเฉลี่ยของผลการประเมินอภิมานรายงานฯ กลุ่มคุณภาพต่ำ
n_1	คือ จำนวนรายงานที่มีคุณภาพสูง
n_2	คือ จำนวนรายงานที่มีคุณภาพต่ำ
S^2	คือ ความแปรปรวนของคะแนนการประเมินอภิมาน

5. วิเคราะห์หาค่าความเที่ยงแบบสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient)

5.1 ออกแบบการวัด 2 Facets ดังนี้ Facet 1 คือ จำนวนตัวบ่งชี้ในแบบประเมิน อภิมาน (I) Facet 2 คือ จำนวนนักประเมินอภิมานที่ประเมินรายงานผลการประเมินคุณภาพภายใน แต่ละฉบับ (R:P) โดยการออกแบบการวิเคราะห์ เป็น I X (R:P) Design กล่าวคือ นักประเมินอภิมาน ต่างกันทำการประเมินรายงานผลการประเมินคุณภาพภายในฉบับเดียวกัน โดยใช้แบบประเมิน อภิมานชุดเดียวกัน วิเคราะห์ด้วยโปรแกรม GENOVA Version 3.1 มีขั้นตอนการวิเคราะห์ ดังนี้

ตารางที่ 3.5 การประมาณค่าความแปรปรวนจากแหล่งต่างๆ ของ I X (R:P) Design ด้วยการคำนวณ G-STUDY

Effect	Variance	Expected Mean Square
I	$[MS(I)-MS(IP)-MS(IR:P)]/n_p n_{R:P}$	$n_p n_{R:P} \sigma_I^2 + n_{R:P} \sigma_{IP}^2 + \sigma_{IR:P}^2$
P	$[MS(P)-MS(IP)-MS(IR:P)]/n_i n_{R:P}$	$n_i n_{R:P} \sigma_P^2 + n_{R:P} \sigma_{IP}^2 + \sigma_{IR:P}^2$
R:P	$[MS(R:P)-MS(IR:P)]/n_i$	$n_i \sigma_{R:P}^2 + \sigma_{IR:P}^2$
IP	$[MS(IP)-MS(IR:P)]/n_{R:P}$	$n_{R:P} \sigma_{IP}^2 + \sigma_{IR:P}^2$
IR:P	$[MS(IR:P)]$	$\sigma_{IR:P}^2$

เมื่อ	I	แทน	จำนวนตัวบ่งชี้ในแบบประเมินอภิมาน
	P	แทน	จำนวนรายงานผลการประเมินคุณภาพภายใน
	R:P	แทน	จำนวนนักประเมินอภิมานที่ประเมินรายงานผลการประเมินคุณภาพภายในแต่ละฉบับ
	IP	แทน	ผลร่วมระหว่างคะแนนประเมินอภิมานกับรายงานผลการประเมินคุณภาพภายใน
	IR:P	แทน	ส่วนที่เหลือ

5.2 การคำนวณ D-Study เพื่อคำนวณหาค่าสัมประสิทธิ์การสรุปอ้างอิง (G-Coefficient) โดยใช้สูตร (Shavelson, and Webb, 1991) ดังนี้

สัมประสิทธิ์การสรุปอ้างอิงเชิงสัมพัทธ์ (Relative G-Coefficient)

$$\rho_{Rel}^2 = \left(\frac{\sigma_P^2}{\sigma_P^2 + \sigma_{Rel}^2} \right)$$

สัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์ (Absolute G-Coefficient)

$$\rho_{Abs}^2 = \left(\frac{\sigma_P^2}{\sigma_P^2 + \sigma_{Abs}^2} \right)$$

บทที่ 4

ผลการวิเคราะห์ข้อมูล

การวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อพัฒนามาตรฐานการประเมินอภิमान สำหรับประเมินรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาในประเทศไทย 2) เพื่อตรวจสอบคุณภาพของมาตรฐานการประเมินอภิमान ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา 3) เพื่อประยุกต์ใช้มาตรฐานการประเมินอภิमानในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา ระดับอุดมศึกษา

ในการนำเสนอผลการวิเคราะห์ข้อมูลและแปลความหมายผลการวิเคราะห์ข้อมูลในครั้งนี้ เพื่อให้การนำเสนอผลการวิเคราะห์ข้อมูลมีความสะดวกและเข้าใจตรงกัน ผู้วิจัยจึงได้กำหนดสัญลักษณ์และความหมายสำหรับใช้แทนค่าสถิติและตัวแปรในการนำเสนอต่างๆ ดังนี้

สัญลักษณ์ที่ใช้แทนค่าสถิติในการวิจัย

N	หมายถึง	จำนวนตัวอย่างในการวิจัย
M	หมายถึง	ค่าเฉลี่ยเลขคณิต
SD	หมายถึง	ค่าความเบี่ยงเบนมาตรฐาน (standard deviation)
Min	หมายถึง	ค่าต่ำสุด
Max	หมายถึง	ค่าสูงสุด
r	หมายถึง	ค่าความสัมพันธ์ของผลการประเมินระหว่างนักประเมินอภิमान
t	หมายถึง	ค่าความแตกต่างของค่าเฉลี่ยของผลการประเมินระหว่างผู้ทรงคุณวุฒิ กับนักประเมินอภิमान
p-value	หมายถึง	ระดับนัยสำคัญทางสถิติ
df	หมายถึง	องศาอิสระ
EMS	หมายถึง	การประมาณค่าความแปรปรวน
$\hat{\sigma}_{(i)}^2$	หมายถึง	ความแปรปรวนของตัวบ่งชี้ในแบบประเมินอภิमान
$\hat{\sigma}_{(p)}^2$	หมายถึง	ความแปรปรวนของรายงานผลการประเมินคุณภาพภายใน
$\hat{\sigma}_{(r.p)}^2$	หมายถึง	ความแปรปรวนของนักประเมินอภิमानที่ประเมินรายงานผลการประเมินฯ

$\hat{\sigma}_{(ip)}^2$	หมายถึง	ความแปรปรวนของผลร่วมระหว่างตัวบ่งชี้ในแบบประเมินอภิमानกับ รายงานผลการประเมินคุณภาพภายใน
$\hat{\sigma}_{(ir;p)}^2$	หมายถึง	ความแปรปรวนของส่วนที่เหลือ
$\hat{\sigma}_{\delta}^2$	หมายถึง	ความแปรปรวนเชิงสัมพัทธ์
$\hat{\sigma}_{\Delta}^2$	หมายถึง	ความแปรปรวนเชิงสัมบูรณ์
$\hat{\rho}_{\delta}^2$	หมายถึง	สัมประสิทธิ์การสรุปอ้างอิงเชิงสัมพัทธ์
$\hat{\rho}_{\Delta}^2$	หมายถึง	สัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์
χ^2	หมายถึง	ค่าไคสแควร์
GFI	หมายถึง	ค่าดัชนีวัดระดับความกลมกลืน
AGFI	หมายถึง	ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว
RMSEA	หมายถึง	ค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ

ผลการวิเคราะห์

ผู้วิจัยจึงขอนำเสนอผลการวิเคราะห์ข้อมูลแยกเป็น 4 ตอน ดังนี้ ตอนที่ 1 การพัฒนาและวิเคราะห์มาตรฐานการประเมินอภิमान ตอนที่ 2 ผลการอบรมพัฒนานักประเมินอภิमान ตอนที่ 3 ผลการตรวจสอบคุณภาพของเครื่องมือประเมินอภิमान และตอนที่ 4 ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา มีรายละเอียดดังต่อไปนี้

ตอนที่ 1 การพัฒนา และวิเคราะห์มาตรฐานการประเมินอภิमान

มาตรฐานการประเมินอภิमानประกอบด้วย แบ่งออกเป็น 3 ส่วน คือ 1. มาตรฐานการประเมิน 2. การสังเคราะห์มาตรฐานการประเมินอภิमानที่พัฒนาขึ้น และ 3. การตรวจสอบความตรงเชิงโครงสร้างของมาตรฐานการประเมินอภิमान โดยมีรายละเอียดดังต่อไปนี้

1. มาตรฐานการประเมินอภิमान

มาตรฐานการประเมินอภิमान ประกอบด้วย 5 มาตรฐาน และ 38 ตัวบ่งชี้ ดังมีรายละเอียดดังต่อไปนี้

มาตรฐานที่ 1 ความตรงของการประเมิน (Validity)

คุณลักษณะของงานประเมิน ที่มีการจัดการเอกสารในการประเมิน การวิเคราะห์บริบท กำหนดวัตถุประสงค์และออกแบบการประเมิน การเก็บรวบรวมสารสนเทศ วิเคราะห์ แปลผล

และสรุปสารสนเทศอย่างมีคุณภาพ และตรวจสอบการประเมินอย่างถูกต้องสมเหตุสมผล ประกอบด้วย 13 ตัวบ่งชี้ คือ

- Va1.1 การระบุบริบทของสถาบันที่ส่งผลกระทบต่อการศึกษา
- Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน
- Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมินอย่างชัดเจน
- Va1.4 การออกแบบการประเมินที่สอดคล้องกับสภาพ และบริบทของสถาบัน
- Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสารในการประเมิน
- Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ
- Va1.7 การตรวจสอบความถูกต้องของข้อมูล
- Va1.8 การตรวจสอบคุณภาพของสารสนเทศ ตรงกับความต้องการของผู้ใช้ข้อมูล
- Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูลอย่างถูกต้องตามวิธีการที่กำหนด
- Va1.10 การแปลผลและสรุปผลการประเมินถูกต้องตามหลักฐานที่ปรากฏ
- Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ
- Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง
- Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมินจากผู้ที่เกี่ยวข้อง

ตัวบ่งชี้การประเมินในมาตรฐานความตรงของการประเมิน มีรายละเอียดดังต่อไปนี้

- Va1.1 การระบุบริบทของสถาบันที่ส่งผลกระทบต่อการศึกษา

คำอธิบาย การแสดงถึงประวัติ ภูมิหลัง ข้อมูลพื้นฐานของสถาบัน การดำเนินงาน และผลการดำเนินงาน รวมถึงสภาพแวดล้อมของสถาบันที่มีอิทธิพลและการวิเคราะห์ผลกระทบต่อการศึกษา ได้แก่ นโยบายการจัดการศึกษา แผนยุทธศาสตร์อุดมศึกษาระยะยาว 15 ปี สภาพสถาบัน สภาพแวดล้อมทางเศรษฐกิจ สังคม การเมือง คุณลักษณะข้อมูลที่เกี่ยวข้องกับระบบประเมินคุณภาพ สภาพการแข่งขันของรัฐและเอกชน เป็นต้น

- Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน

คำอธิบาย อัตลักษณ์ของสถาบัน เป็นลักษณะเด่นของสถาบัน ที่เป็นที่ยอมรับของสาธารณชน โดยสถาบันมีการกำหนดนโยบาย และแผนการดำเนินงาน เพื่อให้สามารถบรรลุจุดเน้นหรืออัตลักษณ์ของสถาบันได้

Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมินอย่างชัดเจน

คำอธิบาย วัตถุประสงค์ในการประเมิน เป็นการแสดงจุดมุ่งหมายของการประเมินอย่างชัดเจน เพื่อให้ผู้มีส่วนได้ส่วนเสียสามารถตรวจสอบได้ว่า การประเมินได้ครอบคลุมนโยบาย การพัฒนาการประเมิน รวมถึงข้อตกลงเบื้องต้นของปัจจัยภายนอกที่มีส่วนในการวางแผนการประเมินในสามารถบรรลุวัตถุประสงค์ได้

Va1.4 การออกแบบการประเมินที่สอดคล้องกับสภาพ และบริบทของสถาบัน

คำอธิบาย การออกแบบการประเมิน เป็นการบรรยายลักษณะของการประเมินที่น่าเชื่อถือ การแสดงหลักฐานสนับสนุนคุณภาพการประเมิน ตั้งแต่วิธีการเก็บรวบรวมข้อมูล ได้แก่ การเลือกกลุ่มตัวอย่าง เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ชนิดของข้อมูล ตลอดจนเทคนิคการวิเคราะห์ข้อมูล ช่วงเวลาและทรัพยากรที่จำเป็น โดยมีการพิจารณาถึงผลกระทบต่อข้อค้นพบและข้อสรุปในการประเมิน และข้อจำกัดในการประเมิน นอกจากนี้มีการออกแบบการเขียนรายงานการประเมินที่สอดคล้องกับสภาพและบริบทของสถาบัน

Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสารในการประเมิน

คำอธิบาย เอกสารในการประเมิน ได้แก่ รายงานการประเมินตนเอง (SAR) เอกสารอ้างอิง ได้แก่ แผนงาน โครงการ/กิจกรรม รายงานผลการดำเนินงาน และหลักสูตร การจัดการเรียนการสอน การประเมินผล รวมถึงผลงานต่างๆ เป็นต้น มีการแสดงรายละเอียดในการพิจารณา วิเคราะห์ ความแตกต่างของเอกสาร หลักฐานจากแหล่งต่างกัน

Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ

คำอธิบาย ข้อมูลที่ใช้ในการประเมินคุณภาพภายใน ต้องมีความน่าเชื่อถือ ได้มาจากการแหล่งการเก็บรวบรวมข้อมูลจริงที่ได้มาตรฐาน จากหลายแหล่ง ทั้งการสังเกตสถานที่ การดำเนินงาน การสัมภาษณ์ ผู้ที่เกี่ยวข้องทุกฝ่าย และการศึกษาจากเอกสารหลักฐานที่เกี่ยวข้อง และการรวบรวมข้อมูลอื่นๆ โดยนำมาเปรียบเทียบเพื่อพิจารณาความสอดคล้อง น่าเชื่อถือของข้อมูล

Va1.7 การตรวจสอบความถูกต้องของข้อมูล

คำอธิบาย ความถูกต้องของข้อมูลในการประเมินคุณภาพภายใน นอกจากข้อมูลการประเมินมาจากหลายแหล่งแล้ว ยังต้องมีการตรวจสอบว่าข้อมูลเหล่านั้นเป็นจริง โดยทำการตรวจสอบซ้ำจากการบันทึกข้อมูล หรือเปรียบเทียบข้อมูลก่อนและหลังการประเมิน เพื่อแสดงหลักฐานที่ยืนยันได้ว่า ข้อมูลมีความสอดคล้องกัน และมีการควบคุมความคลาดเคลื่อนจากปัจจัย

ที่ส่งผลต่อคุณภาพของข้อมูล เพื่อความถูกต้องในการตัดสินใจผลการประเมินตามมาตรฐาน และเกณฑ์การประเมิน

Va1.8 การตรวจสอบคุณภาพของสารสนเทศ ตรงกับความต้องการของผู้ใช้ข้อมูล

คำอธิบาย สารสนเทศในการประเมิน มาจากข้อมูลที่รวบรวมมา มีความพอเพียง และผ่านการวิเคราะห์แล้ว ซึ่งมีความถูกต้อง สมบูรณ์ และน่าเชื่อถือ นอกจากนี้ยังสอดคล้องกับความต้องการของผู้ใช้ข้อมูล

Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูลอย่างถูกต้องตามวิธีการที่กำหนด

คำอธิบาย ตัวอย่างชี้ในการประเมินคุณภาพภายใน มีทั้งตัวอย่างชี้เชิงปริมาณ และเชิงคุณภาพ ซึ่งต้องมีกระบวนการวิเคราะห์ข้อมูลที่ต้อง ตรวจสอบความถูกต้องได้จากการรวบรวมข้อมูลดิบ ใน common dataset ของสถาบัน และสูตรในการคำนวณในคู่มือการประกันคุณภาพการศึกษา ภายในสถานศึกษา ระดับอุดมศึกษาของสกอ. และเปรียบเทียบกับมาตรฐาน เกณฑ์การประเมินคุณภาพภายใน

Va1.10 การแปลผลและสรุปผลการประเมินถูกต้องตามหลักฐานที่ปรากฏ

คำอธิบาย การแปลผลและสรุปผลการประเมิน โดยการตัดสินใจผลการประเมิน ดีความจากข้อค้นพบ เป็นการตัดสินใจคุณค่าของข้อค้นพบเทียบกับมาตรฐาน เกณฑ์การประเมิน จากข้อมูลเชิงประจักษ์ (empirical finding) ตามตัวอย่างชี้ และสรุปผลการตัดสินใจผลการประเมินจากการพิจารณาร่วมกันของทีมผู้ประเมิน

Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ

คำอธิบาย การเขียนรายงานผลการประเมิน ตามข้อค้นพบจากการประเมินและผลการวิเคราะห์ข้อมูล และการสรุปผลการประเมิน ซึ่งความน่าเชื่อถือของข้อสรุปมาจากหลักฐานอ้างอิงและความเชี่ยวชาญของทีมผู้ประเมิน ที่แสดงข้อค้นพบทั้งเชิงบวก และเชิงลบที่ได้จากการประเมิน มีการนำเสนอทั้งจุดเด่น และจุดที่ควรพัฒนาในการประเมินอย่างชัดเจนและสมบูรณ์

Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง

คำอธิบาย การรายงานผลการประเมิน มาจากข้อค้นพบ ข้อสรุปจากการประเมิน ที่ปราศจากการบิดเบือนผลการประเมิน โดยตัวอย่างชี้ในการประเมินมีข้อมูลสนับสนุนเหตุผลในการวิเคราะห์ผลการประเมิน และแสดงให้เห็นว่า มีการพิจารณาตัดสินใจผลการประเมินจากการมีส่วนร่วมของผู้มีส่วนได้ส่วนเสียทุกกลุ่ม

Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมินจากผู้ที่เกี่ยวข้อง

คำอธิบาย ผู้ที่เกี่ยวข้องหรือผู้มีส่วนได้ส่วนเสียหลักสามารถเข้ามามีส่วนร่วมในการทบทวนร่างรายงานการประเมินเพื่อตรวจสอบความสอดคล้องของข้อเท็จจริง ข้อผิดพลาดหรือข้อยกเว้น ที่มีผลต่อการตีความข้อค้นพบจากการประเมินที่อาจทำให้เกิดการพิจารณาที่ไม่ถูกต้อง ผู้ประเมินจะสามารถยอมรับการเปลี่ยนแปลงเพื่อแก้ไขให้ถูกต้อง แต่ทั้งนี้ต้องยึดหลักการความเป็นอิสระของผู้ประเมินที่จะสามารถสรุปผลการประเมินจากข้อค้นพบด้วยตนเอง

มาตรฐานที่ 2 อรรถประโยชน์ (Utility)

คุณลักษณะของงานประเมิน ที่เป็นประโยชน์ต่อผู้มีส่วนได้ส่วนเสีย และผู้ที่เกี่ยวข้อง มีการตัดสินคุณค่า และรายงานผลการประเมินอย่างชัดเจน มีการเผยแพร่ผลการประเมินทันความต้องการของผู้ใช้ผลการประเมิน และเป็นประโยชน์ในการวางแผน ปรับปรุงการดำเนินงานในทิศทางที่เสนอแนะ ประกอบด้วย 10 ตัวบ่งชี้ คือ

Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องในการประเมิน

Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม

Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล โดยคำนึงถึงข้อจำกัดของทรัพยากรการประเมิน

Ut 2.4 การตรวจสอบการตัดสินผลการประเมินสอดคล้องกับสภาพจริง

Ut 2.5 ผลการประเมินเป็นประโยชน์ ซึ่งให้เห็นทิศทางการปรับปรุง การพัฒนาการดำเนินงานของสถาบัน

Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด

Ut 2.7 รายงานผลการประเมินครบถ้วน ครอบคลุมผลการประเมิน และข้อเสนอแนะเพื่อการพัฒนา

Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ และมีสื่อประกอบ

Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา

Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม

ตัวบ่งชี้การประเมินในมาตรฐานอรรถประโยชน์ มีรายละเอียดดังต่อไปนี้

Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องในการประเมิน

คำอธิบาย ผู้มีส่วนได้ส่วนเสียในการประเมิน หมายถึง ผู้ที่เกี่ยวข้องกับการประเมินและได้รับผลกระทบจากการประเมิน ได้แก่ ผู้บริหาร อาจารย์ บุคลากร นักศึกษา ศิษย์เก่า ผู้ใช้บัณฑิต ผู้ปกครอง และชุมชน เป็นต้น

Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม

คำอธิบาย รายงานผลการประเมินมีการระบุถึงช่วงเวลา ระยะเวลาในการเก็บรวบรวมข้อมูลประเมิน โดยมีกำหนดการประเมินที่ระบุรายละเอียดการเก็บรวบรวมข้อมูล ผู้ให้ข้อมูล และมีการประสานงานกับผู้ให้ข้อมูล ในการจัดเตรียมข้อมูลให้สอดคล้องกับการประเมิน

Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล โดยคำนึงถึงข้อจำกัดของทรัพยากรการประเมิน

คำอธิบาย เทคนิคการเก็บรวบรวมข้อมูล ได้แก่ การเลือกผู้ตัวอย่าง เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ชนิดของข้อมูล โดยวางแผนการประเมินเพื่อรวบรวมสารสนเทศที่จำเป็น ซึ่งพิจารณาจากทรัพยากรที่จำเป็น ผลกระทบต่อข้อค้นพบและข้อสรุปในการประเมิน รวมทั้งข้อจำกัดในการประเมิน

Ut 2.4 การตรวจสอบการตัดสินผลการประเมินสอดคล้องกับสภาพจริง

คำอธิบาย สารสนเทศในรายงานมีความสมบูรณ์ โครงสร้างครบถ้วน นำเสนอชัดเจน โดยการพิจารณาถึงความเหมาะสมในการตีความผลการประเมินที่มีหลักฐานที่น่าเชื่อถือเพียงพอต่อการตัดสินคุณค่า ตามมาตรฐานการประเมินที่กำหนด และสอดคล้องกับสภาพจริง

Ut 2.5 ผลการประเมินเป็นประโยชน์ ซึ่งเห็นทิศทางการปรับปรุง การพัฒนาการดำเนินงานของสถาบัน

คำอธิบาย ผลการประเมินและข้อสรุป สอดคล้องกับวัตถุประสงค์ของการประเมินและการตัดสินโดยอิงจากการประเมิน โดยต้องมีความชัดเจน และสามารถนำไปใช้ประโยชน์ในการวางแผนดำเนินการ ซึ่งเห็นทิศทางการพัฒนาที่ปฏิบัติได้จริงทันเหตุการณ์

Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด

คำอธิบาย รายงานผลการประเมินตามรูปแบบที่สำนักงานคณะกรรมการการอุดมศึกษากำหนด ประกอบด้วย รายงานคณะผู้ประเมิน บทนำ การวางแผนและการประเมิน วิธีการตรวจสอบความถูกต้องน่าเชื่อถือของข้อมูล ผลการประเมินรายตัวบ่งชี้ องค์ประกอบ

คุณภาพ มาตรฐานการอุดมศึกษา และมุมมองด้านการบริหารจัดการ มีจุดแข็ง จุดอ่อน และข้อเสนอแนะของแต่ละองค์ประกอบและมาตรฐาน รวมถึงภาคผนวก

Ut 2.7 รายงานผลการประเมินครบถ้วน ครอบคลุมผลการประเมิน และข้อเสนอแนะเพื่อการพัฒนา

คำอธิบาย รายงานผลการประเมินมีการนำเสนอสารสนเทศที่สำคัญในรายงานการประเมินครบถ้วน ครอบคลุมผลการประเมิน โดยแสดงจุดแข็งและจุดอ่อนของข้อสรุปในการประเมิน ซึ่งข้อสรุปในการประเมินอิงจากข้อค้นพบ ที่ตอบคำถามในการประเมินตามข้อมูลที่สืบค้นได้ โดยระบุปัญหาและข้อเสนอแนะที่สามารถนำไปปฏิบัติได้ และปรับปรุงการประเมินครั้งต่อไป

Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ และมีสื่อประกอบ

คำอธิบาย รายงานผลการประเมินจะต้องมีความชัดเจน เขียนด้วยภาษาเชิงวิชาการที่อ่านเข้าใจง่าย และบอกถึงข้อจำกัดของการประเมิน โดยมีสื่อประกอบด้วย รูปภาพ ตารางและตัวอย่างที่ง่ายต่อความเข้าใจ

Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา

คำอธิบาย รายงานผลการประเมินต้องมีความครบถ้วน สมบูรณ์ จัดทำรายงานเสร็จทันตามกำหนด และตรงกับช่วงเวลาที่ใช้ข้อมูลต้องการนำไปพิจารณาวางแผน ปรับปรุงการจัดการดำเนินการในองค์กร และการประเมิน เป็นไปตามกฎกระทรวงว่าด้วย ระบบ หลักเกณฑ์ และวิธีการประกันคุณภาพการศึกษาภายในสถานศึกษา ระดับอุดมศึกษา พ.ศ. 2546

Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม

คำอธิบาย มีการเผยแพร่ผลการประเมินในช่วงเวลาที่เหมาะสมตามความต้องการของผู้ใช้ผลการประเมิน และเปิดเผยผลการประเมินอย่างกว้างขวาง ใช้ภาษาเชิงวิชาการที่ครอบคลุมผู้ใช้ผลการประเมินทุกกลุ่ม

มาตรฐานที่ 3 จริยธรรมในการประเมิน (Ethicality)

คุณลักษณะของงานประเมินที่มีการกำหนดวิธีการประเมินไว้เหมาะสมกับสภาพจริง ได้รับการยอมรับ มีการคำนึงถึงผู้ที่เกี่ยวข้องหลายกลุ่ม มีการปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงสิทธิมนุษยชน และประโยชน์สาธารณะ ทำการประเมินที่มีความถูกต้อง ยุติธรรมต่อผู้รับผลการประเมิน รวมถึงเปิดเผยผลการประเมินแก่ผู้ที่ได้รับผลกระทบจากการประเมิน ประกอบด้วย 9 ตัวบ่งชี้ คือ

Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมินตลอดช่วงเวลาประเมิน

Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง และปราศจากอคติการประเมิน

Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงความต้องการ
จำเป็นของผู้รับการประเมิน

Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน

Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัดที่มีผลกระทบต่อผลการประเมิน

Et 3.6 การกำหนดวิธีการประเมินโดยคำนึงถึงความปลอดภัย รักษาความลับ
และไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล

Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง โดยคำนึงถึงความแตกต่าง หลากหลายของผู้ที่
เกี่ยวข้อง

Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย

Et 3.9 การประเมินที่ครบถ้วน ชัดเจนสอดคล้องกับมาตรฐานและ ตัวบ่งชี้

ตัวบ่งชี้การประเมินในมาตรฐานจริยธรรมในการประเมิน มีรายละเอียดดังต่อไปนี้

Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมินตลอดช่วงเวลาประเมิน

คำอธิบาย การจัดการประเมิน ที่มีคณะผู้ประเมินที่มีศักยภาพในการประเมิน
สามารถกำหนดวิธีการดำเนินงานและตารางการปฏิบัติงานประเมินให้เหมาะสมตามสภาพจริง
โดยมีการเจรจากับผู้ที่เกี่ยวข้องที่มีส่วนร่วมในการประเมินตลอดช่วงเวลาในการทำการประเมิน

Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง และปราศจากอคติการประเมิน

คำอธิบาย การประเมินโดยคำนึงถึงความแตกต่างของผู้ที่เกี่ยวข้องหลากหลาย
กลุ่ม ปราศจากอคติ และการตัดสินใจที่อาจขัดขวางการประเมิน ผลการประเมินเป็นที่ยอมรับ
จากสถาบัน โดยผู้ประเมินเสนอผลการประเมินตามวัตถุประสงค์ของการประเมิน ที่สะท้อนให้เห็น
จุดเด่น/จุดที่ควรพัฒนา/ทิศทางการพัฒนาในอนาคต

Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงความต้องการ
จำเป็นของผู้รับการประเมิน

คำอธิบาย การประเมินคุณภาพ มีการพิจารณานำผลการประเมินมาปรับปรุง
คุณภาพการประเมิน โดยเปรียบเทียบผลการประเมินกับความต้องการจำเป็นของผู้รับผลการ
ประเมิน เพื่อให้ได้รับประโยชน์สูงสุดจากการประเมินอย่างต่อเนื่อง

Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน

คำอธิบาย มีการทำเอกสารการประเมิน กำหนดการประเมินและวิธีการประเมิน
ที่ประยุกต์ใช้เฉพาะกลุ่มที่รับการประเมินและผู้มีส่วนได้ส่วนเสียอื่นที่เกี่ยวข้องอย่างเป็นลาย

ลักษณะอักษรระหว่างผู้ประเมินกับสถาบัน ตามวัตถุประสงค์การประเมิน กระบวนการประเมิน ทรัพยากรในการประเมิน

Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัดที่มีผลกระทบต่อผลการประเมิน

คำอธิบาย การประเมินมีการเปิดเผยผลการประเมิน ข้อจำกัดในการประเมินให้กลุ่มบุคคลทุกฝ่ายที่ได้รับผลกระทบ หรือบุคคลอื่นที่เกี่ยวข้องได้รับทราบผลการประเมิน โดยรายงานผลการประเมินมีรายละเอียดที่ครอบคลุม ตอบสนองความต้องการของผู้รับผลการประเมิน

Et 3.6 การกำหนดวิธีการประเมินโดยคำนึงถึงความปลอดภัย รักษาความลับ และไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล

คำอธิบาย การกำหนดขอบเขตและวิธีการประเมินการประเมินที่คำนึงถึงความปลอดภัย คุณค่า ความปลอดภัย รักษาความลับ และพิจารณาสิทธิมนุษยชนของผู้รับการประเมิน และผู้มีส่วนได้ส่วนเสียอื่น ๆ ในการออกแบบและความแตกต่างของสารสนเทศในการประเมิน โดยไม่ก่อให้เกิดความเสียหายและความทุกข์ใจที่จะเกิดกับผู้มีส่วนร่วมในการประเมิน

Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง โดยคำนึงถึงความแตกต่าง หลากหลายของผู้ที่เกี่ยวข้อง

คำอธิบาย การจัดการประเมินที่คำนึงถึงความสัมพันธ์กับผู้ที่เกี่ยวข้องในการประเมินที่มีความแตกต่าง และหลากหลาย โดยมีการศึกษาถึงภูมิหลังของบุคคลที่เกี่ยวข้องกับการประเมิน เพื่อสร้างปฏิสัมพันธ์ที่ดี

Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย

คำอธิบาย การดำเนินการประเมินที่ให้ความเป็นธรรมกับผู้ที่เกี่ยวข้องกับการประเมิน โดยมีการตรวจสอบและบันทึกจุดเด่น จุดที่ควรพัฒนาของสถาบันที่ทำการประเมินเพื่อให้ข้อเสนอแนะที่มีความเป็นกลางต่อผู้ที่นำผลการประเมินไปใช้ประโยชน์

Et 3.9 การประเมินที่ครบถ้วน ชัดเจนสอดคล้องกับมาตรฐานและ ตัวบ่งชี้

คำอธิบาย การตัดสินผลการประเมินต้องมาจากฐานข้อมูลที่ถูกต้อง ครบถ้วน ทั้งการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การลงข้อสรุปข้อมูล การรายงานผลการประเมิน สอดคล้องกับมาตรฐานในการประเมิน

มาตรฐานที่ 4 ความน่าเชื่อถือของนักประเมิน (Credibility)

คุณลักษณะของงานประเมินที่ได้มาจากนักประเมินที่มีสมรรถนะในการประเมิน ได้รับการยอมรับ และไม่มีผลประโยชน์ทับซ้อนเพื่อสร้างความน่าเชื่อถือของข้อค้นพบและสารสนเทศที่ได้รับจากการประเมิน ประกอบด้วย 4 ตัวบ่งชี้ คือ

Cr4.1 ผู้ประเมินมีความสามารถ ในการประเมิน

Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน

Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมินอย่างเหมาะสม

Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อนของผู้ประเมิน

ตัวบ่งชี้ในมาตรฐานความน่าเชื่อถือของนักประเมิน มีรายละเอียดดังต่อไปนี้

Cr4.1 ผู้ประเมินมีความสามารถ ในการประเมิน

คำอธิบาย ผู้ประเมินต้องมีสมรรถนะในการประเมิน ได้แก่ มีความรู้ ความสามารถ ทักษะ และความชำนาญหรือประสบการณ์ในการประเมินที่เหมาะสม โดยสามารถวัด วิเคราะห์ ข้อมูล แปลผล และเผยแพร่ข้อค้นพบอย่างชัดเจน มีเหตุผลเพียงพอสำหรับการตัดสินใจผลการประเมิน เขียนรายงานผลการประเมินได้ชัดเจน นอกจากนี้ใช้เทคนิคการดำเนินการทั้งเชิงคุณภาพและเชิง ปริมาณ

Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน

คำอธิบาย ผู้ประเมินมีทักษะระหว่างบุคคล ได้แก่ ทักษะการสื่อสารทางการ เขียน การนำเสนอผลงาน และทักษะต่างวัฒนธรรม ทักษะแก้ปัญหาความขัดแย้ง ส่งเสริมให้เกิด ปฏิสัมพันธ์ระหว่างบุคคล มีการบันทึกจุดแข็งและข้อจำกัดของการประเมิน รวมถึงทักษะการ นำเสนอข้อเสนอแนะเพื่อพัฒนาที่สามารถนำไปปฏิบัติได้

Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมินอย่างเหมาะสม

คำอธิบาย ผู้ประเมินมีความสามารถในการจัดการประเมินตั้งแต่เริ่มต้นจนกระทั่ง เสร็จสิ้น โดยสามารถสร้างกรอบแนวทางการประเมินที่สอดคล้องกับบริบทของสถาบัน มีการ ออกแบบการประเมินอย่างรัดกุม โดยประสานงานเกี่ยวกับงบประมาณ การใช้ทรัพยากรร่วมกันเช่น ข้อมูล ความเชี่ยวชาญ บุคลากร เครื่องมือปฏิบัติงาน มีการใช้เทคโนโลยีให้เหมาะสม และจัดการ ประเมินให้เหมาะสมกับกับช่วงระยะเวลาที่กำหนด มีการเสนอทิศทางการพัฒนาสถาบันตามจุดเน้น เพื่อให้บรรลุเป้าหมาย

Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อนของผู้ประเมิน

คำอธิบาย ผู้ประเมินต้องคำนึงถึงผลประโยชน์ที่เกี่ยวข้องกับสาธารณะและสวัสดิภาพของสังคมเป็นหลัก โดยมีการเปิดเผยผลประโยชน์ทับซ้อนที่เกิดขึ้นอย่างซื่อสัตย์ เพื่อไม่ให้เกิดการประนีประนอมหรือละเว้นในกระบวนการประเมินและผลการประเมินที่มีอิทธิพลจากกลุ่มบุคคล ที่มีความขัดแย้งระหว่างกลุ่มที่เกี่ยวข้องกับการประเมิน

มาตรฐานที่ 5 ประสิทธิภาพต้นทุน (Cost-effectiveness)

คุณลักษณะของงานประเมิน ที่คำนึงถึงความคุ้มค่าของทรัพยากรที่จำเป็นในการประเมิน มีค่าใช้จ่ายในการประเมินอย่างสมเหตุสมผล มีหลักฐานชัดเจนและมีความโปร่งใส ตรวจสอบได้ของต้นทุนการประเมิน ประกอบด้วย 2 ตัวบ่งชี้ คือ

Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า

Ce5.2 การจัดสรรงบประมาณสำหรับการประเมินอย่างเหมาะสม

ตัวบ่งชี้ในมาตรฐานประสิทธิภาพต้นทุน มีรายละเอียดดังต่อไปนี้

Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า

คำอธิบาย การประเมินที่ให้สารสนเทศที่มีคุณค่าเพียงพอสำหรับตัดสินใจผลการประเมิน สามารถประเมินได้บรรลุเป้าหมาย โดยมีการจัดการทรัพยากรในการประเมิน พิจารณาจากการใช้ทรัพยากรอย่างคุ้มค่า ได้แก่ ค่าใช้จ่ายที่เป็นเงิน เวลา บุคลากร สถานที่ในการจัดการประเมิน โดยผู้รับการประเมินมีความพึงพอใจ

Ce5.2 การจัดสรรงบประมาณสำหรับการประเมินอย่างเหมาะสม

คำอธิบาย การจัดสรรงบประมาณเพื่อเป็นค่าใช้จ่ายในการประเมินอย่างสมเหตุสมผล มีหลักฐานแสดงชัดเจน ได้แก่ มีจัดทำโครงการ หรือจัดตั้งงบประมาณสำหรับการประเมินร่วมกับ สถาบันที่รับการประเมิน และระบุรายละเอียดค่าใช้จ่ายในการประเมินสามารถตรวจสอบได้ โดยมีการแสดงผลเกี่ยวกับต้นทุน และข้อมูลที่เป็นในการประเมิน

2. การสังเคราะห์มาตรฐานการประเมินอภิมานที่พัฒนาขึ้น

การสังเคราะห์มาตรฐานการประเมินอภิมานที่พัฒนาขึ้น กับมาตรฐานการประเมินใน 7 แนวทางประเมิน ได้แก่ 1. Program Evaluation Standards (1994), 2. The African Evaluation guidelines (2002), 3. The DeGEval's Standards (2003), 4. SEVAL's Evaluation Standards (2004), 5. Guiding Principles for Evaluators (2003), 6. Essential Skills Series in Evaluation (1999), 7. Essential Competencies for Program Evaluators (2006) และการประเมิน

อภิมานของ สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) (สมศ.) พบว่า มีความสอดคล้องกับมาตรฐานการประเมินอภิมานใน 5 มาตรฐาน 38 ตัวบ่งชี้ โดยมีตัวบ่งชี้ที่สอดคล้องกับ program evaluation standard ทั้งหมด 30 ตัวบ่งชี้ โดยอีก 8 ตัวบ่งชี้ สอดคล้องกับมาตรฐานอื่นๆ ได้แก่ ตัวบ่งชี้ Va1.4 สอดคล้องกับมาตรฐานการประเมินของ AEA, CES, ECPE และการประเมินอภิมานของ สมศ. ตัวบ่งชี้ Va1.7 สอดคล้องกับมาตรฐานการประเมินของ CES, ECPE และการประเมินอภิมานของ สมศ. ตัวบ่งชี้ Ut2.3 สอดคล้องกับมาตรฐานการประเมินของ AEA, CES, ECPE ตัวบ่งชี้ Va1.11 และ Et3.9 สอดคล้องกับมาตรฐานการประเมินของ CES ตัวบ่งชี้ Va1.2 และ Ut2.6 สอดคล้องกับการประเมินอภิมานของสมศ. ส่วนตัวบ่งชี้ Ut2.9 สอดคล้องตามกฎกระทรวง การรายงานผลการประกันคุณภาพการศึกษา พ.ศ. 2546 ของสำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) ดังตารางที่ 4.1

ตารางที่ 4.1 สังกะระเห้มาตรฐานการประเมินอภิมาน

มาตรฐาน/ตัวบ่งชี้	1 PES	2 AEG	3 DeGEval	4 SEVAL	5 AEA	6 CES	7 ECPE	8 สมศ	Meta- evaluation STD
มาตรฐานที่ 1 มาตรฐานความตรงของการประเมิน									
Va1.1 การระบุบริบทของสถาบันที่ส่งผลกระทบต่อการประเมิน	A2	A2	A2	A2			Situational Analysis		✓
Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน								4.2	✓
Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมินอย่างชัดเจน	A3	A3	U2, A3	U2, A3		Improving Program		1.1	✓
Va1.4 การออกแบบการประเมินที่สอดคล้องกับสภาพ และบริบทของสถาบัน					Honesty	Understanding	Systematic Inquiry	1.2	✓
Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสารในการประเมิน	A1	A1	A1	A1		Improving Program	Systematic Inquiry		✓
Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	A4	A4	A4	A4	Systematic Inquiry		Systematic Inquiry	2.1	✓
Va1.7 การตรวจสอบความถูกต้องของข้อมูล						Improving Program	Systematic Inquiry	2.1	✓
Va1.8 การตรวจสอบคุณภาพของสารสนเทศ และตรงกับความต้องการของผู้ใช้ข้อมูล	A5, A6	A5, A6	A5	A5		Evaluating for Results	Systematic Inquiry	2.2	✓
Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูลอย่างถูกต้องตามวิธีการที่กำหนด	A8, A9	A8, A9	A7	A7	Systematic Inquiry	Evaluating for Results	Systematic Inquiry	2.3	✓
Va1.10 การแปลผลและสรุปผลการประเมินถูกต้องตามหลักฐานที่ปรากฏ	A7, A10	A7, A10	A6, A8	A6, A8	Systematic Inquiry		Systematic Inquiry	2.4	✓
Va1.11 การเขียนรายงานผล ตามข้อค้นพบ ทั้งเชิงบวกและเชิงลบ						Understanding			✓
Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง	A11	A11	P4	A9			Systematic Inquiry		✓

ตารางที่ 4.1 (ต่อ)

มาตรฐาน/ตัวบ่งชี้	1 PES	2 AEG	3 DeGEval	4 SEVAL	5 AEA	6 CES	7 ECPE	8 สมศ	Meta- evaluation STD
Va1.13 การตรวจสอบความถูกต้องของผล การประเมินจากผู้ที่เกี่ยวข้อง	A12	A12	A9	A10				3.1	✓
มาตรฐานที่ 2 อรรถประโยชน์ (Utility)									
U2.1 การระบุผู้มีส่วนได้ส่วนเสียที่ เกี่ยวข้องในการประเมิน	U1	U1	U1	U1	Honesty	Evaluation Framework	Situational Analysis		✓
U2.2 การระบุช่วงเวลาและระยะเวลาใน การเก็บรวบรวมข้อมูลที่เหมาะสม	U3	U3	U4	U4					✓
U2.3 การกำหนดเทคนิคในการเก็บ รวบรวมข้อมูล โดยคำนึงถึงข้อจำกัดของ ทรัพยากรการประเมิน					Honesty	Understanding	Project Management		✓
U2.4 การตรวจสอบการตัดสินใจผลการ ประเมินสอดคล้องกับสภาพจริง	U4	U4	U5	U5	Systematic Inquiry	Understanding	Systematic Inquiry	3.1	✓
U2.5 ผลการประเมินเป็นประโยชน์ ชี้ให้เห็นทิศทางการปรับปรุง การ พัฒนาการดำเนินงานของสถาบัน	U7	U7	U8	U8		Understanding	Systematic Inquiry	3.3	✓
U2.6 การเขียนรายงานสอดคล้องกับ รูปแบบที่กำหนด								3.2	✓
U2.7 รายงานผลการประเมินครบถ้วน ครอบคลุมผลการประเมิน และ ข้อเสนอแนะเพื่อการพัฒนา	U5	U5	U6	U6		Evaluation Framework		4.1	✓
U2.8 รายงานผลการประเมินเขียนด้วย ภาษาเชิงวิชาการ และมีสื่อประกอบ	U5	U5	U6	U6		Understanding	Interpersonal		✓
U2.9 การจัดทำรายงานผลการประเมิน เสร็จทันเวลา	ตามกฎกระทรวง การรายงานผลการประกันคุณภาพการศึกษา พ.ศ. 2546 ของ สกอ.								✓
U2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ ผลการประเมินทุกกลุ่ม	U6	U6	U7	U7		Evaluating for Results			✓
มาตรฐานที่ 3 จริยธรรมในการประเมิน (Ethicality)									
E3.1 การสื่อสารกับผู้มีส่วนร่วมในการ ประเมินตลอดช่วงเวลาประเมิน	F1	F1	F1	F1	Honesty		Project Management		✓
E3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ เกี่ยวข้อง และปราศจากอคติการประเมิน	F2	F2	F2	F2	Responsibilities		Situational Analysis	3.3	✓
E3.3 การปรับปรุงคุณภาพการประเมินอย่าง ต่อเนื่อง โดยคำนึงถึงความต้องการจำเป็น ของผู้รับการประเมิน	P1	P1	-	-	Responsibilities	Improving Program			✓
E3.4 การจัดทำเอกสารข้อตกลงการ ประเมิน	P2	P2	P1	P1	Respect for People		Project Management		✓

ตารางที่ 4.1 (ต่อ)

มาตรฐาน/ตัวบ่งชี้	1 PES	2 AEG	3 DeGEval	4 SEVAL	5 AEA	6 CES	7 ECPE	8 สมศ.	Meta- evaluation STD
Ei3.5 การเปิดเผยผลการประเมินและ ข้อจำกัดที่มีผลกระทบต่อ การประเมิน	P6	P6	P5	P5	Responsibilities		Situational Analysis		✓
Ei3.6 การกำหนดวิธีการประเมินโดย คำนึงถึงความปลอดภัย รักษาความลับ และ ไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล	P3	P3	P2	P2	Respect for People	Evaluation Framework	Profession		✓
Ei3.7 การติดต่อกับผู้ที่เกี่ยวข้อง โดย คำนึงถึงความแตกต่าง หลากหลายของผู้ ที่เกี่ยวข้อง	P4	P4	-	P3			Interpersonal		✓
Ei3.8 การประเมินที่เป็นธรรมต่อผู้ ที่เกี่ยวข้องทุกฝ่าย	P5	P5	P3	P4		Understanding	Profession		✓
Ei3.9 การประเมินที่ครบถ้วน ชัดเจน สอดคล้องกับมาตรฐานและ ตัวบ่งชี้						Evaluation Framework			✓
มาตรฐานที่ 4 ความน่าเชื่อถือของนักประเมิน (Credibility)									
Cr4.1 ผู้ประเมินมีความสามารถ ในการ ประเมิน	U2	U2	U3	U3	Competence		Profession		✓
Cr4.2 ผู้ประเมินมีทักษะในการสื่อสารผล การประเมิน	U2	U2	U3	U3	Competence		Profession		✓
Cr4.3 ผู้ประเมินมีความสามารถในการ จัดการการประเมินอย่างเหมาะสม	U2	U2	U3	U3	Competence		Profession		✓
Cr4.4 กระบวนการประเมินที่คำนึง ผลประโยชน์ทับซ้อนของผู้ประเมิน	P7	P7	-	P6	Responsibilities				✓
มาตรฐานที่ 5 ประสิทธิภาพต้นทุน (Cost-effectiveness)									
Ce5.1 การจัดการทรัพยากรการประเมิน อย่างคุ้มค่า	F3	F3	F3	F3	Honesty	Improving Program	Project Management		✓
Ce5.2 การจัดสรรงบประมาณสำหรับการ ประเมินอย่างเหมาะสม	P8	P8	-	-		Evaluating for Results	Project Management		✓

หมายเหตุ 1. PES (Program Evaluation Standards, 1994), 2. AEG (The African Evaluation guidelines, 2002), 3. DeGEval (The DeGEval's Standards, 2003), 4. SEVAL (SEVAL's Evaluation Standards, 2004), 5. AEA (Guiding Principles for Evaluators, 2004), 6. CES (Essential Skills Series in Evaluation, 1999), 7. ECPE (Essential Competencies for Program Evaluators, 2006), 8. ประเมิน
อภิมาน ของ สมศ. (2550)

3. การตรวจสอบความตรงเชิงโครงสร้างของมาตรฐานการประเมินอภิมาน

การตรวจสอบความตรงเชิงโครงสร้าง(Construct validity) ของตัวบ่งชี้การประเมิน
อภิมานที่ประกอบด้วย มาตรฐานย่อย 5 มาตรฐาน จากผลการประเมินอภิมานรายงานผลการ
ประเมินคุณภาพภายใน จำนวน 50 แห่ง ซึ่งการวิเคราะห์องค์ประกอบแบ่งออกเป็น 6 ขั้นตอน คือ 1)
การวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานความตรงของการประเมิน 2) การวิเคราะห์

องค์ประกอบเชิงยืนยันของมาตรฐานอรรถประโยชน์ 3) การวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานจริยธรรมในการประเมิน 4) การวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานความน่าเชื่อถือของนักประเมิน 5) การวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานประสิทธิผลต้นทุน และ 6) การวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานการประเมินอภิमान ดังมีรายละเอียดต่อไปนี้

3.1 การตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของตัวบ่งชี้การประเมินอภิमानในมาตรฐานความตรงของการประเมิน (validity) โดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม LISREL 8.80 ผลการวิเคราะห์ของแต่ละตัวบ่งชี้ที่มีรายละเอียดดังนี้

มาตรฐานความตรงของการประเมิน วัดได้จากตัวบ่งชี้ 13 ตัวบ่งชี้ ผลการวิเคราะห์เมทริกซ์สหสัมพันธ์ของตัวบ่งชี้ได้ 51 คู่ มีค่าแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 31 คู่ และระดับ .05 จำนวน 20 คู่ ค่าสัมประสิทธิ์สหสัมพันธ์ต่ำสุดมีค่าเท่ากับ 0.042 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ การระบุบริบทของสถาบัน (Va1.1) การกำหนดแหล่งข้อมูล (Va1.6) และค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดมีค่าเท่ากับ 0.695 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ การวิเคราะห์ข้อมูลและประมวลผลข้อมูล (Va1.9) การแปลผลและสรุปผลการประเมิน (Va1.10) ค่า Bartlett's test of Sphericity Chi-Square มีค่าเท่ากับ 244.542, $df=78$, $p=0.000$ ซึ่งแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 สอดคล้องกับค่าดัชนีรวม Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.727 ซึ่งมีค่ามากกว่า 0.5 แสดงว่าเมทริกซ์สหสัมพันธ์ของตัวบ่งชี้ในมาตรฐานความตรงของการประเมินไม่ใช่เมทริกซ์เอกลักษณะ สามารถนำมาวิเคราะห์องค์ประกอบได้

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า มาตรฐานความตรงของการประเมิน (VALIDITY) มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาจากค่าไคสแควร์ที่แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ($\chi^2= 62.87$ $df= 62$, $p=0.445$) ค่าดัชนีวัดระดับความกลมกลืน (GFI) มีค่าเท่ากับ 0.84 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) มีค่าเท่ากับ 0.76 และค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ (RMSEA) มีค่าเท่ากับ 0.017

ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบของตัวบ่งชี้ แสดงให้เห็นว่า ค่าน้ำหนักองค์ประกอบทุกตัวบ่งชี้มีค่าเป็นบวก ขนาดใกล้เคียงกัน มีค่าตั้งแต่ 0.37 ถึง 0.80 โดยตัวบ่งชี้ที่มีน้ำหนักความสำคัญมากที่สุด คือ การตรวจสอบความถูกต้องของผลการประเมิน (Va1.13) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.80 และมีการแปรผันร่วมกับความตรงของการประเมินร้อยละ 48 รองลงมาคือ การออกแบบการประเมิน (Va1.4) การวิเคราะห์แหล่งความแตกต่างของเอกสาร (Va1.5) การบรรยายถึงวัตถุประสงค์ในการประเมิน (Va1.3) การวิเคราะห์ข้อมูลและประมวลผล

ข้อมูล (Va1.9) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.70, 0.63, 0.55 และ 0.55 ตามลำดับ และมีการแปรผันร่วมกับความตรงของการประเมินร้อยละ 30, 39, 31 และ 30 ตามลำดับ ส่วนตัวบ่งชี้ที่มีน้ำหนักความสำคัญน้อยที่สุด คือ การเขียนรายงานผล ตามข้อค้นพบ (Va1.11) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.37 และมีการแปรผันร่วมกับความตรงของการประเมินร้อยละ 41 แสดงให้เห็นว่า ตัวบ่งชี้เหล่านี้มีความสำคัญในการวัดความตรงของการประเมิน ผลวิเคราะห์แสดงดังตารางที่ 4.2, 4.3 และภาพที่ 4.1

ตารางที่ 4.2 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน

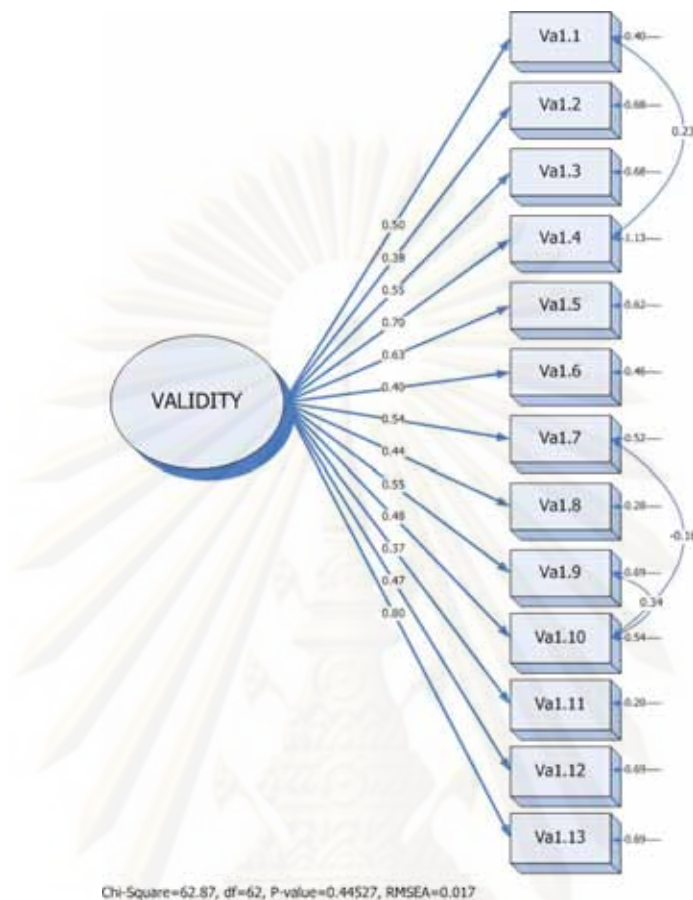
มาตรฐานความตรงของการประเมิน

ตัวบ่งชี้	Va1.1	Va1.2	Va1.3	Va1.4	Va1.5	Va1.6	Va1.7	Va1.8	Va1.9	Va1.10	Va1.11	Va1.12	Va1.13
Va1.1	1.000												
Va1.2	0.367**	1.000											
Va1.3	0.221	0.474**	1.000										
Va1.4	0.162	0.330*	0.309*	1.000									
Va1.5	0.260	0.427**	0.568**	0.251	1.000								
Va1.6	0.042	0.162	0.177	0.291*	0.350*	1.000							
Va1.7	0.432**	0.473**	0.222	0.472**	0.413**	0.256	1.000						
Va1.8	0.184	0.338*	0.488**	0.338*	0.426**	0.410**	0.446**	1.000					
Va1.9	0.199	0.257	0.205	0.399**	0.401**	0.290*	0.313*	0.345*	1.000				
Va1.10	0.110	0.274	0.264	0.415**	0.424**	0.174	0.120	0.289	0.695**	1.000			
Va1.11	0.267	0.300*	0.439**	0.321*	0.308*	0.504**	0.236	0.448**	0.450**	0.403**	1.000		
Va1.12	0.358*	0.153	0.231	0.321*	0.243	0.169	0.194	0.351*	0.320*	0.321	0.441**	1.000	
Va1.13	0.303*	0.327*	0.345*	0.535**	0.440**	0.487**	0.354**	0.360**	0.289*	0.383**	0.470**	0.358**	1.000
M	2.560	2.280	2.340	2.920	2.740	2.840	2.740	2.120	2.900	2.740	3.540	2.780	2.840
SD	0.907	0.991	1.272	1.007	0.803	0.792	0.899	0.689	0.995	0.876	0.579	0.954	1.149

Bartlett's test of Sphericity Chi-Square = 244.542, df=78, p=0.000

Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.727

* p< .05 ,** p<.01



ภาพที่ 4.1 โมเดลมาตรฐานความตรงของการประเมิน

ตารางที่ 4.3 ผลการวิเคราะห์หองค์ประกอบเชิงยืนยันของมาตรฐานความตรงของการประเมิน

ตัวบ่งชี้	น้ำหนักองค์ประกอบ	SE	t	R ²	SC	FS
Va1.1	0.50	0.11	4.53	0.39	0.62	0.14
Va1.2	0.38	0.13	2.87	0.18	0.42	0.07
Va1.3	0.55	0.14	3.99	0.31	0.56	0.11
Va1.4	0.70	0.18	3.90	0.30	0.55	0.05
Va1.5	0.63	0.14	4.54	0.39	0.62	0.13
Va1.6	0.40	0.11	3.59	0.26	0.51	0.12
Va1.7	0.54	0.12	4.30	0.36	0.60	0.18
Va1.8	0.44	0.09	4.71	0.41	0.64	0.21
Va1.9	0.55	0.14	3.89	0.30	0.55	0.03
Va1.10	0.48	0.12	3.84	0.30	0.55	0.15
Va1.11	0.37	0.08	4.75	0.41	0.64	0.25
Va1.12	0.47	0.14	3.44	0.24	0.49	0.09
Va1.13	0.80	0.15	5.23	0.48	0.69	0.15

$\chi^2 = 62.87$ df= 62, p=0.445, GFI=0.84 , AGFI=0.76 , RMSEA=0.017

หมายเหตุ ค่า SC= completely standardized solution, ค่า FS =factor scores regressions

3.2 การตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของตัวบ่งชี้การประเมินปริมาณในมาตรฐานอรรถประโยชน์ (utility) โดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม LISREL 8.80 ผลการวิเคราะห์ของแต่ละตัวบ่งชี้มีรายละเอียดดังนี้

มาตรฐานอรรถประโยชน์ วัดได้จากตัวบ่งชี้ 10 ตัวบ่งชี้ ผลการวิเคราะห์เมทริกซ์สหสัมพันธ์ของตัวบ่งชี้ได้ 27 คู่ มีค่าแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 19 คู่ และระดับ .05 จำนวน 8 คู่ ค่าสัมประสิทธิ์สหสัมพันธ์ต่ำสุดมีค่าเท่ากับ -0.052 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล (Ut2.3) การเขียนรายงานถูกต้อง (Ut2.7) และค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดมีค่าเท่ากับ 0.594 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้การระบุผู้มีส่วนได้ส่วนเสีย (Ut2.1) การเผยแพร่ผลการประเมิน (Ut2.10) ค่า Bartlett's test of Sphericity Chi-Square มีค่าเท่ากับ 173.948, $df=45$, $p=0.000$ ซึ่งแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าดัชนีรวม Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.708 ซึ่งมีค่ามากกว่า 0.5 แสดงว่าเมทริกซ์สหสัมพันธ์ของตัวบ่งชี้ในมาตรฐานอรรถประโยชน์ไม่ใช่เมทริกซ์เอกลักษณะ สามารถนำมาวิเคราะห์องค์ประกอบได้

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า มาตรฐานอรรถประโยชน์ (UTILITY) มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาจากค่าไคสแควร์ที่แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ($\chi^2= 31.62$ $df= 30$, $p=0.38557$) ค่าดัชนีวัดระดับความกลมกลืน (GFI) มีค่าเท่ากับ 0.89 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) มีค่าเท่ากับ 0.79 และค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ (RMSEA) มีค่าเท่ากับ 0.033

ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบของตัวบ่งชี้ แสดงให้เห็นว่า ค่าน้ำหนักองค์ประกอบทุกตัวบ่งชี้มีค่าเป็นบวก มีค่าตั้งแต่ 0.18 ถึง 0.81 โดยตัวบ่งชี้ที่มีน้ำหนักความสำคัญมากที่สุด คือ ผลการประเมินเป็นประโยชน์ (Ut2.5) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.81 และมีการแปรผันร่วมกับอรรถประโยชน์ร้อยละ 50 รองลงมาคือ การเขียนรายงานสอดคล้องกับรูปแบบ (Ut2.6) การระบุผู้มีส่วนได้ส่วนเสีย (Ut2.1) การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล (Ut2.2) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.64, 0.56, และ 0.56 ตามลำดับ และมีการแปรผันร่วมกับอรรถประโยชน์ร้อยละ 46, 29, และ 34 ตามลำดับ ส่วนตัวบ่งชี้ที่มีน้ำหนักความสำคัญน้อยที่สุด คือ การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล (Ut2.3) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.18 และมีการแปรผันร่วมกับอรรถประโยชน์ร้อยละ 10 แสดงให้เห็นว่า ตัวบ่งชี้เหล่านี้มีความสำคัญในการวัดอรรถประโยชน์ ผลวิเคราะห์แสดงดังตารางที่ 4.4, 4.5 และภาพที่ 4.2

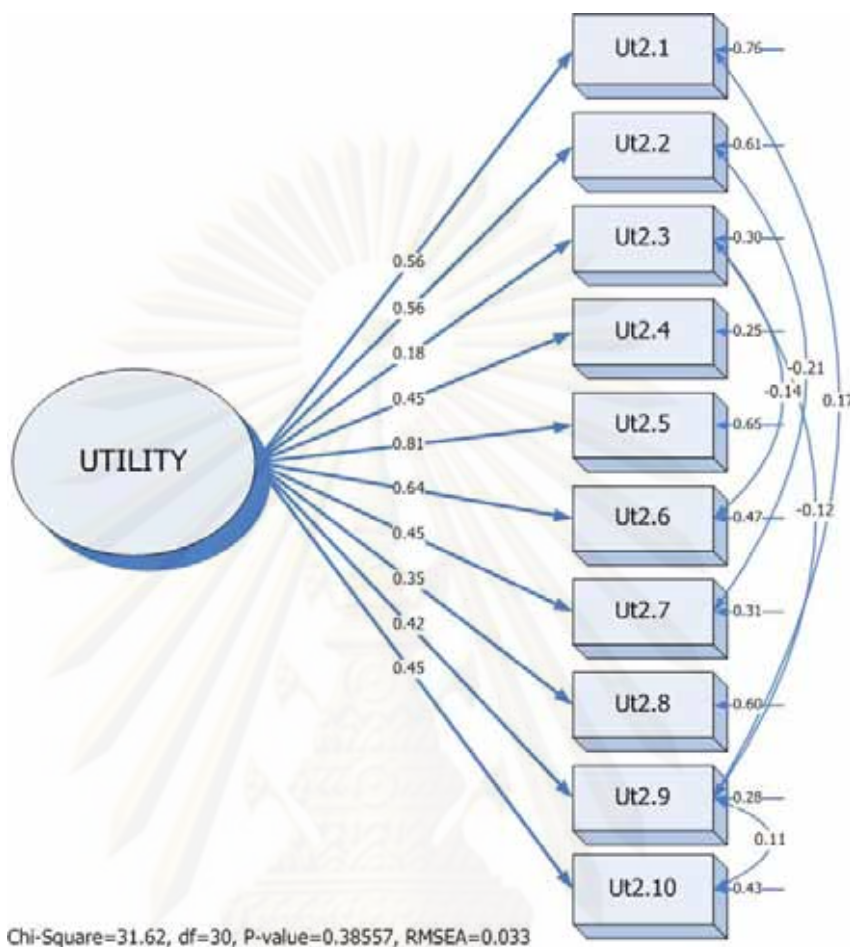
ตารางที่ 4.4 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน
มาตรฐานอรรถประโยชน์

ตัวบ่งชี้	Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6	Ut2.7	Ut2.8	Ut2.9	Ut2.10
Ut2.1	1.000									
Ut2.2	0.302 *	1.000								
Ut2.3	0.259	0.230	1.000							
Ut2.4	0.347 *	0.371 **	0.162	1.000						
Ut2.5	0.417 **	0.350 *	0.242	0.287 *	1.000					
Ut2.6	0.346 *	0.394 **	0.185	0.559 **	0.374 **	1.000				
Ut2.7	0.228	0.263	-0.052	0.518 **	0.283 *	0.576 **	1.000			
Ut2.8	0.440 **	0.061	0.125	0.332 *	0.443 **	0.377 **	0.577 **	1.000		
Ut2.9	0.154	0.424 **	0.162	0.453 **	0.160	0.303	0.136	0.135	1.000	
Ut2.10	0.594 **	0.443 **	-0.070	0.370 **	0.575 **	0.390 **	0.345 *	0.388 **	0.243	1.000
M	3.02	3.18	2.42	3.14	3.22	2.86	3.16	2.88	3.34	3.54
SD	1.040	0.962	0.575	0.670	0.790	1.143	0.934	0.718	0.848	0.676

Bartlett's test of Sphericity Chi-Square = 173.948, df=45, p=0.000

Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.708

* p< .05 ,** p<.01



ภาพที่ 4.2 โมเดลมาตรฐานอรรถประโยชน์

ตารางที่ 4.5 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานอรรถประโยชน์

ตัวบ่งชี้	น้ำหนักองค์ประกอบ	SE	t	R ²	SC	FS
Ut2.1	0.56	0.15	3.84	0.29	0.54	0.03
Ut2.2	0.56	0.14	4.11	0.34	0.58	0.22
Ut2.3	0.18	0.09	2.04	0.10	0.31	0.30
Ut2.4	0.45	0.09	4.98	0.44	0.66	0.21
Ut2.5	0.81	0.15	5.41	0.50	0.71	0.15
Ut2.6	0.64	0.13	5.08	0.46	0.68	0.25
Ut2.7	0.45	0.10	4.55	0.39	0.63	0.32
Ut2.8	0.35	0.12	2.86	0.17	0.41	0.07
Ut2.9	0.42	0.09	4.47	0.38	0.62	0.27
Ut2.10	0.45	0.11	4.04	0.32	0.56	0.06

$\chi^2 = 31.62$ df = 30, p = 0.38557, GFI = 0.89, AGFI = 0.79, RMSEA = 0.033

หมายเหตุ คำ SC = completely standardized solution, คำ FS = factor scores regressions

3.3 การตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของตัวบ่งชี้การประเมินอภิमानในมาตรฐานจริยธรรมในการประเมิน (Ethicality) โดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม LISREL 8.80 ผลการวิเคราะห์ของแต่ละตัวบ่งชี้มีรายละเอียดดังนี้

มาตรฐานจริยธรรมในการประเมิน วัดได้จากตัวบ่งชี้ 9 ตัวบ่งชี้ ผลการวิเคราะห์เมทริกซ์สหสัมพันธ์ของตัวบ่งชี้ได้ 32 คู่ มีค่าแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 25 คู่ และระดับ .05 จำนวน 7 คู่ ค่าสัมประสิทธิ์สหสัมพันธ์ต่ำสุดมีค่าเท่ากับ 0.168 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ การจัดทำเอกสารข้อตกลงการประเมิน (Et3.4) การประเมินที่ครบถ้วน ชัดเจน (Et3.9) และค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดมีค่าเท่ากับ 0.705 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ผลการประเมินเป็นที่ยอมรับ (Et3.2) การประเมินที่เป็นธรรม (Et3.8) ค่า Bartlett's test of Sphericity Chi-Square มีค่าเท่ากับ 178.446, $df=36$, $p=0.000$ ซึ่งแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าดัชนีรวม Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.836 ซึ่งมีค่ามากกว่า 0.5 แสดงว่าเมทริกซ์สัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ในมาตรฐานจริยธรรมในการประเมินไม่ใช่เมทริกซ์เอกลักษณะ สามารถนำมาวิเคราะห์องค์ประกอบได้

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า มาตรฐานจริยธรรมในการประเมิน (ETHICALITY) มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาจากค่าไคสแควร์ที่แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ($\chi^2 = 26.39$ $df = 24$, $p = 0.33374$) ค่าดัชนีวัดระดับความกลมกลืน (GFI) มีค่าเท่ากับ 0.89 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) มีค่าเท่ากับ 0.80 และค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ (RMSEA) มีค่าเท่ากับ 0.045

ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบของตัวบ่งชี้ แสดงให้เห็นว่า ค่าน้ำหนักองค์ประกอบทุกตัวบ่งชี้มีค่าเป็นบวก มีค่าตั้งแต่ 0.35 ถึง 0.64 โดยตัวบ่งชี้ที่มีน้ำหนักความสำคัญมากที่สุด คือ ผลการประเมินเป็นที่ยอมรับ (Et3.2) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.64 และมีการแปรผันร่วมกับจริยธรรมในการประเมิน ร้อยละ 62 รองลงมาคือ การกำหนดวิธีการประเมินโดยคำนึงถึงความปลอดภัยของผู้ให้ข้อมูล (Et3.6) การประเมินที่ครบถ้วน ชัดเจน (Et3.9) การประเมินที่เป็นธรรม (Et3.8) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.63, 0.62, และ 0.57 ตามลำดับ และมีการแปรผันร่วมกับจริยธรรมในการประเมิน ร้อยละ 64, 49, และ 39 ตามลำดับ ส่วนตัวบ่งชี้ที่มีน้ำหนักความสำคัญน้อยที่สุด คือ การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง (Et3.3) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.35 และมีการแปรผันร่วมกับจริยธรรมในการประเมิน ร้อยละ 24 แสดงให้เห็น

ว่า ตัวบ่งชี้เหล่านี้มีความสำคัญในการวัดจริยธรรมในการประเมิน ผลวิเคราะห์แสดง ดังตารางที่ 4.6, 4.7 และภาพที่ 4.3

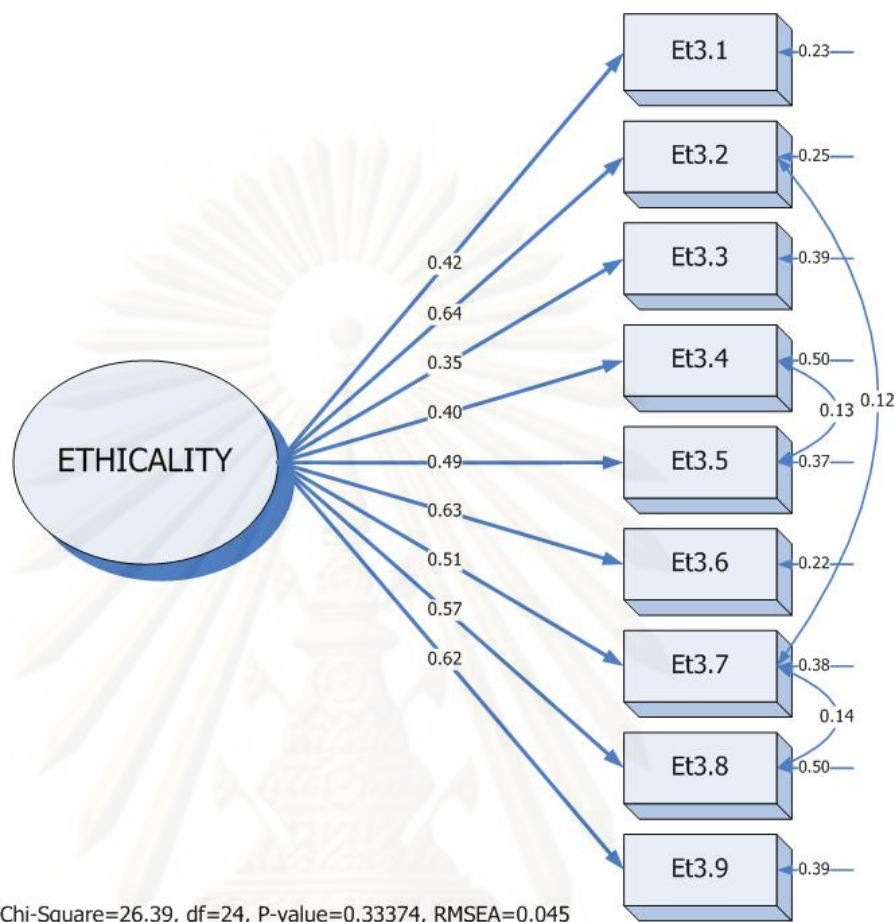
ตารางที่ 4.6 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน
มาตรฐานจริยธรรมในการประเมิน

ตัวบ่งชี้	Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6	Et3.7	Et3.8	Et3.9
Et3.1	1.000								
Et3.2	0.518**	1.000							
Et3.3	0.426**	0.379**	1.000						
Et3.4	0.431**	0.360**	0.317*	1.000					
Et3.5	0.479**	0.540**	0.334*	0.262	1.000				
Et3.6	0.303*	0.571**	0.197	0.511**	0.334*	1.000			
Et3.7	0.551**	0.571**	0.420**	0.464**	0.587**	0.568**	1.000		
Et3.8	0.431**	0.705**	0.338*	0.292*	0.382**	0.460**	0.451**	1.000	
Et3.9	0.302*	0.581**	0.217	0.168	0.524**	0.426**	0.463**	0.625**	1.000
M	3.400	3.540	2.820	2.900	3.040	3.380	2.520	3.300	3.220
SD	0.639	0.813	0.720	0.814	0.880	0.780	0.789	0.814	0.910

Bartlett's test of Sphericity Chi-Square = 178.446, df=36, p=0.000

Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.836

* p< .05 ,** p<.01



ภาพที่ 4.3 โมเดลมาตรฐานจริยธรรมในการประเมิน

ตารางที่ 4.7 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานจริยธรรมในการประเมิน

ตัวบ่งชี้	น้ำหนักองค์ประกอบ	SE	t	R ²	SC	FS
Et3.1	0.42	0.09	4.87	0.43	0.66	0.23
Et3.2	0.64	0.10	6.23	0.62	0.79	0.31
Et3.3	0.35	0.10	3.43	0.24	0.49	0.11
Et3.4	0.40	0.12	3.46	0.25	0.50	0.07
Et3.5	0.49	0.11	4.61	0.39	0.63	0.15
Et3.6	0.63	0.10	6.43	0.64	0.80	0.36
Et3.7	0.51	0.11	4.63	0.40	0.64	0.02
Et3.8	0.57	0.12	4.60	0.39	0.63	0.14
Et3.9	0.62	0.12	5.32	0.49	0.70	0.20

$\chi^2 = 26.39$ df= 24, p=0.33374, GFI=0.89 , AGFI=0.80 , RMSEA=0.045

หมายเหตุ ค่า SC= completely standardized solution, ค่า FS =factor scores regressions

3.4 การตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของตัวบ่งชี้การประเมินอภิमानในมาตรฐานความน่าเชื่อถือของนักประเมิน (Credibility) โดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม LISREL ผลการวิเคราะห์ของแต่ละตัวบ่งชี้มีรายละเอียดดังนี้

มาตรฐานความน่าเชื่อถือของนักประเมิน วัดได้จากตัวบ่งชี้ 4 ตัวบ่งชี้ ผลการวิเคราะห์เมทริกซ์สหสัมพันธ์ของตัวบ่งชี้ได้ 3 คู่ มีค่าแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 2 คู่ และระดับ .05 จำนวน 1 คู่ ค่าสัมประสิทธิ์สหสัมพันธ์ต่ำสุดมีค่าเท่ากับ 0.095 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ ผู้ประเมินมีทักษะในการสื่อสาร (Cr4.2) กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน (Cr4.4) และค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดมีค่าเท่ากับ 0.399 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ผู้ประเมินมีความสามารถในการประเมิน (Cr4.1) ผู้ประเมินมีทักษะในการสื่อสาร (Cr4.2) ค่า Bartlett's test of Sphericity Chi-Square มีค่าเท่ากับ 22.783, $df=6$, $p=0.001$ ซึ่งแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าดัชนีรวม Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.671 ซึ่งมีค่ามากกว่า 0.5 แสดงว่าเมทริกซ์สัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ในมาตรฐานความน่าเชื่อถือของนักประเมินไม่ใช่เมทริกซ์เอกลักษณะ สามารถนำมาวิเคราะห์องค์ประกอบได้

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า มาตรฐานความน่าเชื่อถือของนักประเมิน (CREDIBILITY) มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาจากค่าไคสแควร์ที่แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ($\chi^2= 1.10$ $df= 2$, $p=0.57804$) ค่าดัชนีวัดระดับความกลมกลืน (GFI) มีค่าเท่ากับ 0.99 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) มีค่าเท่ากับ 0.94 และค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ (RMSEA) มีค่าเท่ากับ 0.000

ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบของตัวบ่งชี้ แสดงให้เห็นว่า ค่าน้ำหนักองค์ประกอบทุกตัวบ่งชี้มีค่าเป็นบวก มีค่าตั้งแต่ 0.28 ถึง 0.48 โดยตัวบ่งชี้ที่มีน้ำหนักความสำคัญมากที่สุด คือ ผู้ประเมินมีความสามารถในการประเมิน (Cr4.1) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.48 และมีการแปรผันร่วมกับความน่าเชื่อถือของนักประเมิน ร้อยละ 47 รองลงมาคือ ผู้ประเมินมีความสามารถในการจัดการการประเมิน (Cr4.3) ผู้ประเมินมีทักษะในการสื่อสาร (Cr4.2) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.43 และ 0.39 ตามลำดับ และมีการแปรผันร่วมกับความน่าเชื่อถือของนักประเมิน ร้อยละ 36 และ 31 ตามลำดับ ส่วนตัวบ่งชี้ที่มีน้ำหนักความสำคัญน้อยที่สุด คือ กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน (Cr4.4) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.28 และมีการแปรผันร่วมกับความน่าเชื่อถือของนักประเมิน ร้อยละ 11 แสดงให้เห็นว่า ตัวบ่งชี้เหล่านี้

มีความสำคัญในการวัดความน่าเชื่อถือของนักประเมิน ผลวิเคราะห์แสดงดังตารางที่ 4.8, 4.9 และ ภาพที่ 4.4

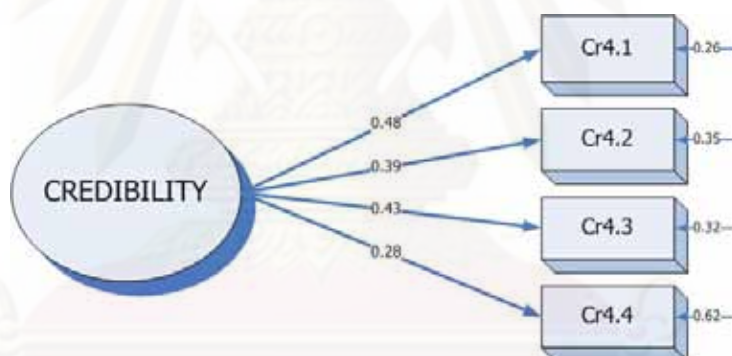
ตารางที่ 4.8 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ใน มาตรฐานความน่าเชื่อถือของนักประเมิน

ตัวบ่งชี้	Cr4.1	Cr4.2	Cr4.3	Cr4.4
Cr4.1	1.000			
Cr4.2	0.399**	1.000		
Cr4.3	0.389**	0.343*	1.000	
Cr4.4	0.246	0.095	0.258	1.000
M	3.14	3.32	3.22	3.14
SD	0.700	0.713	0.708	0.833

Bartlett's test of Sphericity Chi-Square = 22.783, df=6, p=0.001

Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.671

* p< .05 ,** p<.01



Chi-Square=1.10, df=2, P-value=0.57804, RMSEA=0.000

ภาพที่ 4.4 โมเดลมาตรฐานความน่าเชื่อถือของนักประเมิน

ตารางที่ 4.9 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานความน่าเชื่อถือของนักประเมิน

ตัวบ่งชี้	น้ำหนัก องค์ประกอบ	SE	t	R ²	SC	FS
Cr4.1	0.48	0.12	3.85	0.47	0.69	0.61
Cr4.2	0.39	0.12	3.26	0.31	0.55	0.37
Cr4.3	0.43	0.12	3.48	0.36	0.60	0.44
Cr4.4	0.28	0.14	1.96	0.11	0.34	0.15

$\chi^2 = 1.10$ df= 2, p=0.57804, GFI=0.99 , AGFI=0.94, RMSEA=0.000

หมายเหตุ ค่า SC= completely standardized solution, ค่า FS =factor scores regressions

3.5 การตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของตัวบ่งชี้การประเมินอภิमानในมาตรฐานประสิทธิผลต้นทุน (Cost-effectiveness) โดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม LISREL 8.80 ผลการวิเคราะห์ของแต่ละตัวบ่งชี้มีรายละเอียดดังนี้

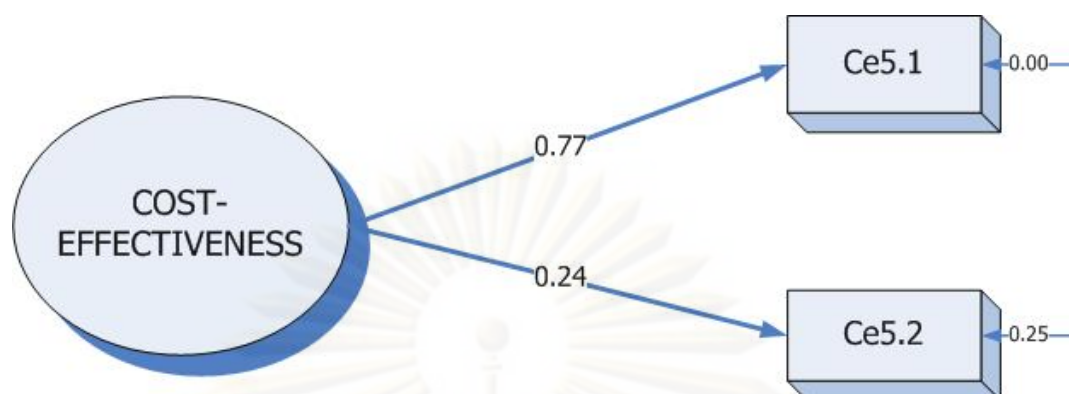
มาตรฐานประสิทธิผลต้นทุน วัดได้จากตัวบ่งชี้ 2 ตัวบ่งชี้ ผลการวิเคราะห์เมทริกซ์สหสัมพันธ์ของตัวบ่งชี้ มีค่าแตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ โดยค่าสัมประสิทธิ์สหสัมพันธ์มีค่าเท่ากับ 0.234 เป็นความสัมพันธ์ระหว่างตัวบ่งชี้ ผู้ประเมินมีทักษะในการสื่อสาร (Ce5.1) กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน (Ce5.2)

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า มาตรฐานประสิทธิผลต้นทุน (COST-EFFECTIVENESS) มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาจากค่าไคสแควร์ที่แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ($\chi^2 = 0.00$ df= 1, p=0.97139) และค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ (RMSEA) มีค่าเท่ากับ 0.000

ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบของตัวบ่งชี้ แสดงให้เห็นว่า ค่าน้ำหนักองค์ประกอบทุกตัวบ่งชี้มีค่าเป็นบวก มีค่าตั้งแต่ 0.24 ถึง 0.77 โดยตัวบ่งชี้ที่มีน้ำหนักความสำคัญมากที่สุด คือ การจัดการทรัพยากรการประเมินอย่างคุ้มค่า (Ce5.1) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.77 และมีการแปรผันร่วมกับประสิทธิผลต้นทุน ร้อยละ 100 ส่วนตัวบ่งชี้ที่มีน้ำหนักความสำคัญน้อยที่สุด คือ การจัดสรรงบประมาณสำหรับการประเมิน (Ce5.2) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.24 และมีการแปรผันร่วมกับประสิทธิผลต้นทุน ร้อยละ 19 แสดงให้เห็นว่าตัวบ่งชี้เหล่านี้มีความสำคัญในการวัดประสิทธิผลต้นทุน ผลวิเคราะห์แสดงดังตารางที่ 4.10, 4.11 และภาพที่ 4.5

ตารางที่ 4.10 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของตัวบ่งชี้ในมาตรฐานประสิทธิผลต้นทุน

ตัวบ่งชี้	Ce5.1	Ce5.2
Ce5.1	1	
Ce5.2	0.234	1
M	2.24	1.34
SD	0.771	0.557



Chi-Square=0.00, df=1, P-value=0.97139, RMSEA=0.000

ภาพที่ 4.5 โมเดลมาตรฐานประสิทธิผลต้นทุน

ตารางที่ 4.11 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของมาตรฐานประสิทธิผลต้นทุน

ตัวบ่งชี้	น้ำหนัก องค์ประกอบ	SE	t	R ²	SC	FS
Ce5.1	0.77	0.08	9.90	1.00	1.00	1.30
Ce5.2	0.24	0.08	3.20	0.19	0.44	-

$\chi^2 = 0.00$ df= 1, p=0.97139, RMSEA=0.000

หมายเหตุ ค่า SC= completely standardized solution, ค่า FS =factor scores regressions

3.6 การตรวจสอบความตรงเชิงโครงสร้าง (construct validity) ของมาตรฐานการประเมินอภิमान (Meta-evaluation) โดยวิธีการวิเคราะห์องค์ประกอบเชิงยืนยันด้วยโปรแกรม LISREL 8.80 โดยใช้คะแนนองค์ประกอบ (factor score) ในการคำนวณ ซึ่งผลการวิเคราะห์ของแต่ละมาตรฐานมีรายละเอียดดังนี้

มาตรฐานการประเมินอภิमान (META-EVALUATION) วัดได้จากมาตรฐาน 5 มาตรฐานย่อย ผลการวิเคราะห์เมตริกซ์สหสัมพันธ์ของมาตรฐานได้ 7 คู่ มีค่าแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 จำนวน 5 คู่ และระดับ .05 จำนวน 2 คู่ ค่าสัมประสิทธิ์สหสัมพันธ์ต่ำสุดมีค่าเท่ากับ 0.127 เป็นความสัมพันธ์ระหว่างมาตรฐานความน่าเชื่อถือของนักประเมิน (Credibility) มาตรฐานประสิทธิภาพการประเมิน (cost-effectiveness) และค่าสัมประสิทธิ์สหสัมพันธ์สูงสุดมีค่าเท่ากับ 0.769 เป็นความสัมพันธ์ระหว่างมาตรฐานความตรงของการประเมิน (validity) กับมาตรฐานอรรถประโยชน์ (utility) ค่า Bartlett's test of Sphericity Chi-Square มีค่าเท่ากับ 100.070, df=10, p=0.000 ซึ่งแตกต่างจากศูนย์อย่างมีนัยสำคัญทางสถิติที่ระดับ .01 ค่าดัชนีรวม Kaiser-Meyer-Olkin Measure of Sampling Adequacy=0.747 ซึ่งมีค่ามากกว่า 0.5

แสดงว่าเมทริกซ์สัมประสิทธิ์สหสัมพันธ์ของมาตรฐานย่อยในมาตรฐานประเมินอภิमानไม่ใช่เมทริกซ์เอกลักษณะ สามารถนำมาวิเคราะห์องค์ประกอบได้

ผลการวิเคราะห์องค์ประกอบเชิงยืนยันพบว่า มาตรฐานการประเมินอภิमान (META-EVALUATION) มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาจากค่าไคสแควร์ที่แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ($\chi^2 = 0.84$, $df = 3$, $p = 0.84067$) ค่าดัชนีวัดระดับความกลมกลืน (GFI) มีค่าเท่ากับ 0.99 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) มีค่าเท่ากับ 0.97 และค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ (RMSEA) มีค่าเท่ากับ 0.000

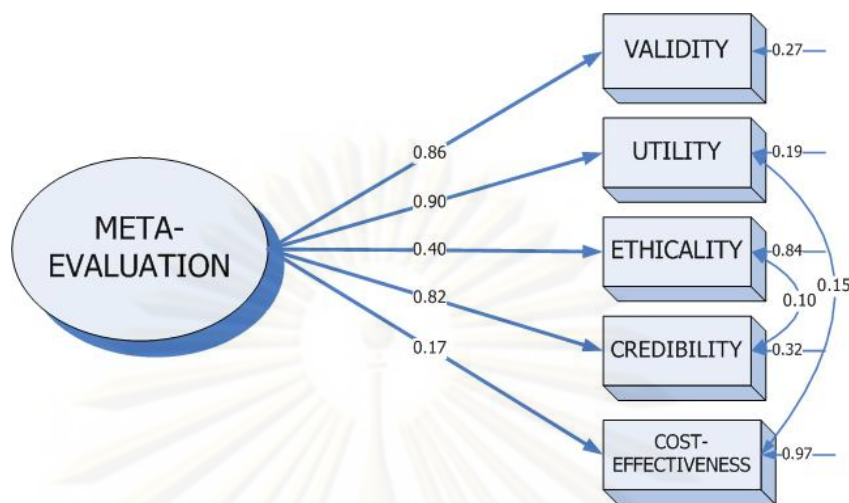
ผลการวิเคราะห์ค่าน้ำหนักองค์ประกอบของมาตรฐานแสดงให้เห็นว่า ค่าน้ำหนักองค์ประกอบทุกมาตรฐานมีค่าเป็นบวก มีค่าตั้งแต่ 0.17 ถึง 0.90 โดยมาตรฐานที่มีน้ำหนักความสำคัญมากที่สุด คือ มาตรฐานอรรถประโยชน์ (utility) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.90 และมีการแปรผันร่วมกับมาตรฐานการประเมินอภิमान ร้อยละ 81 รองลงมาคือ มาตรฐานความตรงของการประเมิน (validity) ความน่าเชื่อถือของนักประเมิน (credibility) และจริยธรรมในการประเมิน (ethicality) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.86, 0.82 และ 0.40 ตามลำดับ และมีการแปรผันร่วมกับมาตรฐานการประเมินอภิमान ร้อยละ 73, 68 และ 16 ตามลำดับ ส่วนมาตรฐานที่มีน้ำหนักความสำคัญน้อยที่สุด คือ ประสิทธิภาพการประเมิน (Cost-effectiveness) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.17 และมีการแปรผันร่วมกับมาตรฐานการประเมินอภิमान ร้อยละ 3 แสดงให้เห็นว่า มาตรฐานเหล่านี้มีความสำคัญในการวัดมาตรฐานการประเมินอภิमान ผลวิเคราะห์แสดงดังตารางที่ 4.12, 4.13 และภาพที่ 4.6

ตารางที่ 4.12 ค่าสัมประสิทธิ์สหสัมพันธ์ของคะแนนมาตรฐานในมาตรฐานการประเมินอภิमान

มาตรฐาน	Validity	Utility	Ethicality	Credibility	Cost-effectiveness
Validity	1.000				
Utility	0.769**	1.000			
Ethicality	0.376**	0.349*	1.000		
Credibility	0.703**	0.741**	0.435**	1.000	
Cost-effectiveness	0.148	0.302*	0.139	0.127	1.000

Bartlett's test of Sphericity Chi-Square = 100.070, $df = 10$, $p = 0.000$
Kaiser-Meyer-Olkin Measure of Sampling Adequacy = 0.747

* $p < .05$, ** $p < .01$



Chi-Square=0.84, df=3, P-value=0.84067, RMSEA=0.000

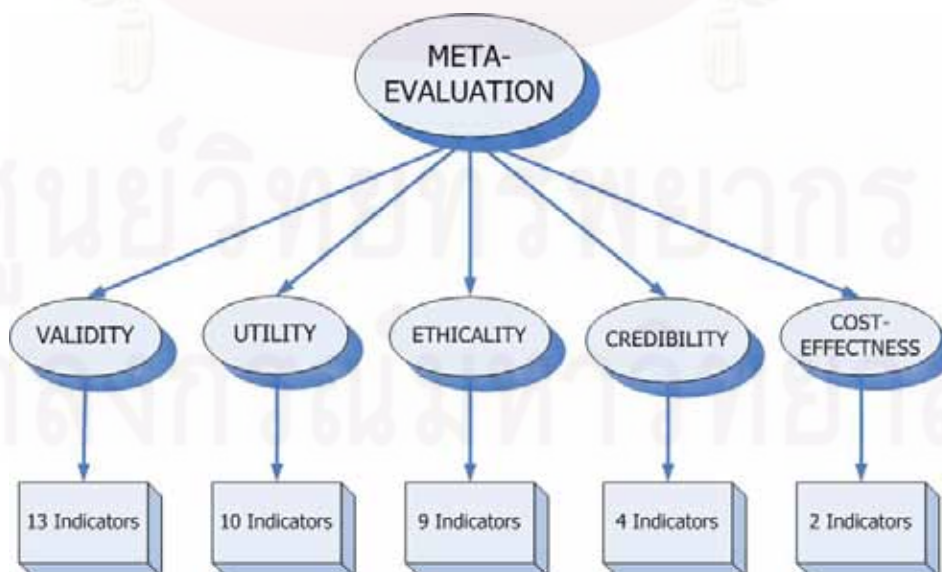
ภาพที่ 4.6 โมเดลมาตรฐานการประเมินอภิมาน

ตารางที่ 4.13 ผลการวิเคราะห์องค์ประกอบเชิงยืนยันของโมเดลการประเมินอภิมาน

มาตรฐาน	น้ำหนักองค์ประกอบ	SE	t	R ²	SC	FS
Validity	0.86	0.12	7.16	0.73	0.86	0.31
Utility	0.90	0.12	7.68	0.81	0.90	0.50
Ethicality	0.40	0.15	2.75	0.16	0.40	0.02
Credibility	0.82	0.12	6.80	0.68	0.82	0.24
Cost-effectiveness	0.17	0.16	1.07	0.03	0.17	-0.06

$\chi^2 = 0.84$ df= 3, p=0.84067, GFI=0.99 , AGFI=0.97 , RMSEA=0.000

หมายเหตุ ค่า SC= completely standardized solution, ค่า FS =factor scores regressions



ภาพที่ 4.7 โมเดลภาพรวมของมาตรฐานการประเมินอภิมาน

ตอนที่ 2 ผลการอบรมพัฒนานักประเมินอภิमान

การอบรมพัฒนานักประเมินอภิमान ตามหลักสูตรฝึกอบรมพัฒนานักประเมินอภิमानที่พัฒนาขึ้น โดยได้จัดฝึกอบรมขึ้น 2 รุ่น คือ **รุ่นที่ 1** จัดขึ้น ณ อาคารประชุมสุข อาชีวอำรุง คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ระหว่างวันที่ 19-20 กันยายน 2552 มีผู้เข้าร่วมอบรม 12 คน และ**รุ่นที่ 2** จัดขึ้น ณ ห้องบงกช อาคาร 10 ชั้น 6 มหาวิทยาลัยศรีปทุม ระหว่างวันที่ 4-5 มกราคม 2553 มีผู้เข้าร่วมอบรม 16 คน ผลการฝึกอบรมนักประเมิน เมื่อเทียบกับวัตถุประสงค์ของหลักสูตร ผู้วิจัยขอเสนอแบ่งออกเป็น 3 ส่วน คือ 1) ผลการฝึกอบรมนักประเมินอภิमान รุ่นที่ 1 2) ผลการฝึกอบรมนักประเมินอภิमान รุ่นที่ 2 และ 3) ผลการวิเคราะห์ความสอดคล้องของผลการประเมินอภิमानของผู้เข้ารับการอบรม ดังมีรายละเอียดต่อไปนี้

2.1 ผลการฝึกอบรมนักประเมินอภิमान รุ่นที่ 1

ผลการฝึกอบรมนักประเมินอภิमानเมื่อเทียบกับวัตถุประสงค์ของหลักสูตร รุ่นที่ 1 ที่จัดขึ้น ณ อาคารประชุมสุข อาชีวอำรุง คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ระหว่างวันที่ 19-20 กันยายน 2552 มีผู้เข้าร่วมอบรม 12 คน พบว่า สอดคล้องเป็นไปตามวัตถุประสงค์หลักสูตร ทั้ง 4 ข้อ คือ 1) เพื่อสร้างความรู้ ความเข้าใจ และบทบาทหน้าที่ของนักประเมินอภิमान 2) เพื่อให้มีความเข้าใจในมาตรฐานและเกณฑ์ในการประเมินอภิमान 3) เพื่อให้มีทักษะในการประเมินอภิमान สามารถตรวจสอบกระบวนการประเมินคุณภาพภายใน และการจัดทำรายงานผลการประเมินคุณภาพภายในได้อย่างถูกต้อง สอดคล้องกับมาตรฐานการประเมินอภิमानและแนวทางการประเมินของสกอ. และ 4) เพื่อให้ผู้ประเมินสามารถนำความรู้ที่ได้รับไปพัฒนาส่งเสริมการประเมินอภิमानภายในหน่วยงาน มีรายละเอียดผลการฝึกอบรมนักประเมินอภิमानตามวัตถุประสงค์ ดังนี้

1) ผู้เข้าร่วมอบรมมีความรู้ ความเข้าใจเกี่ยวกับเนื้อหาการอบรมในส่วนของแนวคิด หลักการการประเมินอภิमानเพิ่มมากขึ้น มีความเข้าใจภาพรวมของบทบาทการเป็นนักประเมินอภิमान แต่ผู้เข้าร่วมอบรมบางคนมีความรู้ในเบื้องต้นจากการอบรมการประเมินอภิमान แต่ยังคงต้องใช้เวลาในการศึกษาเพิ่มเติม สำหรับปัญหาคือ ผู้เข้าร่วมอบรมบางคนยังไม่เข้าใจเกณฑ์ สกอ. เนื่องจากนักประเมินอภิमानต้องเข้าใจการประเมินเช่นเดียวกับผู้ประเมิน หรืออาจต้องมากกว่าผู้ประเมิน

2) ผู้เข้าร่วมอบรมได้ความรู้เพิ่มขึ้น ในการกำหนดเกณฑ์ในแต่ละมาตรฐานและตัวบ่งชี้ มีความเข้าใจเกณฑ์การประเมินเพิ่มมากขึ้น และสามารถนำมาเป็นเกณฑ์ในการประเมินอภิमानได้ แต่ในส่วนของ การปฏิบัติยังไม่ชัดเจนในหลายๆประเด็น ผู้เข้าร่วมอบรมเห็นว่า

ควรปรับคำอธิบาย และนิยามตัวบ่งชี้ และเกณฑ์ในแต่ละระดับให้ชัดเจน เพื่อความเข้าใจตรงกันในการประเมิน โดยเสนอแนะว่าควรปรับปรุงคำที่ระบุในเกณฑ์การให้คะแนน (Rubrics) ให้ชัดเจนสอดคล้องกัน นอกจากนี้ ควรปรับปรุงเอกสารคู่มือการประเมินอภิมานให้ชัดเจนมากขึ้นและควรมีนิยามศัพท์เฉพาะในบางคำที่ต้องตีความให้ตรงกัน รวมถึงหลักฐานอ้างอิงที่ใช้ประกอบการตัดสินใจให้คะแนน

3) ผู้เข้าร่วมอบรมนักประเมินอภิมาน สามารถทำการประเมินอภิมานรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษาทั้ง 5 แห่งได้ โดยแบ่งกลุ่มรับผิดชอบในการประเมิน พิจารณาตามมาตรฐาน ตัวบ่งชี้ และเกณฑ์การประเมินอภิมานที่กำหนดให้ และสามารถเขียนรายงานการประเมินอภิมานสรุปผลการประเมินตามมาตรฐาน และมาตรฐานการประเมินอภิมานได้ รวมทั้งสามารถสรุปจุดแข็ง จุดอ่อนของการประเมิน และให้ข้อเสนอแนะจากการประเมิน ที่เป็นประโยชน์ในการปรับปรุงการประเมินตามแนวทางของสกอ. ต่อไปได้

4) ผู้เข้าร่วมอบรมส่วนใหญ่เห็นว่า ความรู้ที่ได้รับจากการอบรมสามารถนำไปใช้ได้จริงและมีประโยชน์มากโดยนำไปใช้ในการปฏิบัติงานและประยุกต์ใช้ในการประเมินผล การประเมินอภิมานในหน่วยงาน ด้านการประกันคุณภาพ และสามารถนำไปตรวจสอบ ความน่าเชื่อถือของรายงานผลการประเมินคุณภาพภายใน ซึ่งจะทำให้การเขียนรายงานผลการประเมินคุณภาพภายในเป็นมาตรฐานเดียวกัน นอกจากนี้ได้รับความรู้เกี่ยวกับแนวทาง แนวปฏิบัติ ในการประเมินอภิมาน และการพิจารณามาตรฐาน ตัวบ่งชี้ เกณฑ์ รวมถึงได้ฝึกทักษะจากการปฏิบัติ โดยมีการแลกเปลี่ยนประสบการณ์จากผู้เข้าร่วมอบรม ซึ่งความรู้ที่ได้รับมีประโยชน์ เพราะอาจจำเป็นต้องมีการทำการประเมินอภิมานในอนาคต แต่มีผู้เข้าร่วมอบรมบางคน เห็นว่าสามารถนำไปใช้ในการปฏิบัติงานได้ในระดับหนึ่ง แต่ยังไม่ไปปฏิบัติได้ยาก เนื่องจากยังมีปัญหาเกี่ยวกับตัวบ่งชี้ เกณฑ์การให้คะแนน (Rubrics) ในบางมาตรฐานที่ประเมินได้ยาก จำเป็นต้องทดลองใช้และปรับให้เข้ากับบริบทของหน่วยงาน และปัญหาเกี่ยวกับพื้นฐานความรู้ และประสบการณ์ด้านการประเมินคุณภาพภายในของผู้เข้าร่วมอบรมยังน้อย ดังตารางที่ 4.14 ภาพกิจกรรมอบรมดังกล่าว

2.2 ผลการฝึกอบรมนักประเมินอภิมาน รุ่นที่ 2

ผลการฝึกอบรมนักประเมินอภิมานเมื่อเทียบกับวัตถุประสงค์ของหลักสูตร รุ่นที่ 2 ที่จัดขึ้น ณ ห้องบงกช อาคาร 10 ชั้น 6 มหาวิทยาลัยศรีปทุม ระหว่างวันที่ 4-5 มกราคม 2553 มีผู้เข้าร่วมอบรม 16 คน พบว่า สอดคล้องเป็นไปตามวัตถุประสงค์หลักสูตรทั้ง 4 ข้อ คือ 1) เพื่อสร้างความรู้ ความเข้าใจ และบทบาทหน้าที่ของนักประเมินอภิมาน 2) เพื่อให้มีความเข้าใจในมาตรฐานและเกณฑ์ในการประเมินอภิมาน 3) เพื่อให้มีทักษะในการประเมินอภิมาน สามารถ

ตรวจสอบกระบวนการประเมินคุณภาพภายใน และการจัดทำรายงานผลการประเมินคุณภาพภายใน ได้อย่างถูกต้อง สอดคล้องกับมาตรฐานการประเมินอภิมานและแนวทางการประเมินของสกอ. และ4) เพื่อให้ผู้ประเมินสามารถนำความรู้ที่ได้รับไปพัฒนาส่งเสริมการประเมินอภิมานภายในหน่วยงาน มีรายละเอียดผลการฝึกอบรมนักประเมินอภิมานตามวัตถุประสงค์ ดังนี้

1) ผู้เข้าร่วมอบรมมีความรู้ ความเข้าใจเกี่ยวกับเนื้อหาการอบรมในส่วนของแนวคิด หลักการประเมินอภิมานเพิ่มมากขึ้น โดยสามารถอธิบายถึงประโยชน์ของการประเมินอภิมาน การนำการประเมินอภิมานไปประยุกต์ใช้ในการประเมิน การพัฒนาผู้ประเมิน การปรับปรุง การเขียนรายงานการประเมินและSAR รวมถึงประโยชน์ในการปฏิบัติงาน และการพัฒนาการประกันคุณภาพทั้งระบบ การประเมินอภิมานภายในและภายนอก

2) ผู้เข้าร่วมอบรมมีความรู้ ความเข้าใจในมาตรฐาน และเกณฑ์การประเมินอภิมานที่กำหนดให้ โดยสามารถทำการประเมินอภิมานตามเกณฑ์ได้ และมีข้อเสนอแนะเพื่อพิจารณาปรับปรุงเกณฑ์ในการประเมินให้มีความชัดเจนมากยิ่งขึ้น โดยเฉพาะการตีความของแต่ละเกณฑ์ และเกณฑ์การพิจารณาที่ไม่มีกำหนดไว้ในรูปแบบของสำนักงานคณะกรรมการการอุดมศึกษา

3) ผู้เข้าร่วมอบรมนักประเมินอภิมาน สามารถฝึกปฏิบัติการประเมินอภิมาน รายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษาทั้ง 5 แห่งได้ โดยแบ่งกลุ่มรับผิดชอบในการประเมิน พิจารณาตามมาตรฐาน ตัวบ่งชี้ และเกณฑ์การประเมินอภิมานที่กำหนดให้ และสามารถเขียนรายงานการประเมินอภิมานสรุปผลการประเมินตามมาตรฐาน และมาตรฐานการประเมินอภิมานได้ถูกต้อง รวมทั้งสามารถสรุปจุดแข็ง จุดอ่อนของการประเมิน และให้ข้อเสนอแนะจากการประเมิน ที่เป็นประโยชน์ในการปรับปรุงการประเมินตามแนวทางของสกอ. ต่อไปได้

4) ผู้เข้าร่วมอบรมส่วนใหญ่เห็นว่า ความรู้ที่ได้รับจากการอบรมสามารถนำไปใช้ได้จริงและมีประโยชน์มากต่อตัวผู้ประเมิน และการนำผลการประเมินอภิมานไปใช้พัฒนา เพราะทำให้สามารถพัฒนาตัวผู้ประเมินให้มีคุณภาพได้มาตรฐานเดียวกัน อีกทั้งสามารถนำไปปรับปรุงวิธีการเขียนรายงานการประเมินที่ดี รวมถึงการสร้างวัฒนธรรมในการตรวจประเมินให้เป็นมาตรฐานเดียวกัน โดยพัฒนาการประเมินให้มีความตรง น่าเชื่อถือได้รับการยอมรับมากขึ้น และสอดคล้องกับวัตถุประสงค์การประเมิน นอกจากนี้ การประเมินอภิมานเป็นการตรวจสอบยืนยันความถูกต้องของงานประเมิน และความโปร่งใสของกลุ่มผู้ประเมิน สามารถให้ข้อมูลที่ถูกต้องย้อนกลับไปยังสถาบันที่มีการตรวจประเมินได้อีกด้วย ซึ่งเป็นประโยชน์ต่อการประกันคุณภาพทั้ง

ระบบอย่างชัดเจน แต่ยังมีปัญหา/อุปสรรคในการประเมินอภิमान คือ เกณฑ์การประเมินต้องมีความชัดเจน เป็นรูปธรรม ยังมีความแตกต่างของรูปแบบรายงานการประเมิน และความสอดคล้องของรายงานผลการประเมินกับตัวบ่งชี้การประเมินอภิमान ตัวบ่งชี้บางตัวยังมีความซ้ำซ้อนกัน รวมทั้งมาตรฐาน และการยอมรับของผู้ประเมินอภิमानจากสถาบัน ดังตารางที่ 4.14 ภาพกิจกรรมอบรมดังกล่าวภาพผนวก ค

ตารางที่ 4.14 สรุปผลการอบรมนักประเมินอภิमानทั้งสองรุ่น

วัตถุประสงค์ของหลักสูตร	ผลการอบรม รุ่น 1	ผลการอบรม รุ่น 2
1. ความรู้ ความเข้าใจ และ บทบาทหน้าที่ของนักประเมิน อภิमान	เข้าใจแนวคิด หลักการการประเมิน อภิमानเพิ่มมากขึ้น ภาพรวมของ บทบาทการเป็นนักประเมินอภิमान แต่ยังไม่ชัดเจนในหลายๆประเด็น ต้องใช้เวลาในการศึกษาเพิ่มเติม	เข้าใจแนวคิด หลักการการประเมิน อภิमानเพิ่มขึ้น สามารถอธิบายถึง ประโยชน์ของการประเมินอภิमान นำไปประยุกต์ใช้ในการประเมิน การพัฒนาผู้ประเมิน การปรับปรุง การเขียนรายงานการประเมินและ SAR รวมถึงประโยชน์ในการ ปฏิบัติงาน
2. ความเข้าใจในมาตรฐานและ เกณฑ์ในการประเมินอภิमान	เข้าใจการกำหนดเกณฑ์ในแต่ละ มาตรฐาน มาตรฐานและตัวบ่งชี้ สามารถนำมาเป็นเกณฑ์ในการ ประเมินอภิमानได้ แต่ในการปฏิบัติ ยังไม่ชัดเจนในหลายๆประเด็น ควรปรับคำอธิบาย และนิยาม ตัวบ่งชี้ และเกณฑ์ในแต่ละระดับให้ ชัดเจน เพื่อความเข้าใจตรงกันในการประเมิน	สามารถทำการประเมินอภิमान ตามเกณฑ์ได้ และมีข้อเสนอแนะ ปรับปรุงเกณฑ์ในการประเมินให้มีความชัดเจนมากยิ่งขึ้น โดยเฉพาะ การตีความของแต่ละเกณฑ์ และ เกณฑ์การพิจารณาที่ไม่มีกำหนด ไว้ในรูปแบบของสำนักงาน คณะกรรมการการอุดมศึกษา
3. ทักษะในการประเมินอภิमान ตรวจสอบกระบวนการประเมิน คุณภาพภายใน และการจัดทำ รายงานผลการประเมินคุณภาพ ภายในได้อย่างถูกต้อง	สามารถทำการประเมินอภิमान และสามารถเขียนรายงานการ ประเมินอภิमान โดยสรุปผลการ ประเมินตามมาตรฐาน และ มาตรฐานการประเมินอภิमानได้ สามารถสรุปจุดแข็ง จุดอ่อนของ การประเมิน และให้ข้อเสนอแนะ จากการประเมิน ที่เป็นประโยชน์ในการปรับปรุงการประเมินตาม แนวทางของสกอ.	สามารถทำการประเมินอภิमान เขียนรายงานการประเมินอภิमान สามารถสรุปผลการประเมินตาม มาตรฐาน และมาตรฐานการ ประเมินอภิमान โดยสามารถสรุป จุดแข็ง จุดอ่อนของการประเมิน และให้ข้อเสนอแนะจากการ ประเมิน ตามแนวทางของสกอ.

ตารางที่ 4.14 (ต่อ)

วัตถุประสงค์ของหลักสูตร	ผลการอบรม รุ่น 1	ผลการอบรม รุ่น 2
4. ผู้ประเมินสามารถนำความรู้ที่ได้รับไปพัฒนาส่งเสริมการประเมินอภิมานภายใน	สามารถนำไปใช้ประโยชน์ โดยนำไปใช้ในการปฏิบัติงานและการประเมินผล การประเมินอภิมานในหน่วยงาน ด้านการประกันคุณภาพ และตรวจสอบ ความน่าเชื่อถือของรายงานผลการประเมินคุณภาพภายใน ซึ่งจะทำให้การเขียนรายงานผลการประเมินคุณภาพภายในเป็นมาตรฐานเดียวกัน ได้ฝึกทักษะจากการปฏิบัติ โดยมีการแลกเปลี่ยนประสบการณ์จากผู้เข้าร่วมอบรม ซึ่งความรู้ที่ได้รับมีประโยชน์ เพราะอนาคต จำเป็นต้องมีการทำการประเมินอภิมาน แต่ยังไม่ไปปฏิบัติได้ยาก	สามารถนำไปใช้ประโยชน์ต่อตัวผู้ประเมิน และการนำผลการประเมินอภิมานไปใช้พัฒนา สามารถพัฒนาตัวผู้ประเมินให้มีคุณภาพได้ มาตรฐานเดียวกัน อีกทั้งสามารถปรับปรุงวิธีการเขียนรายงานการประเมินที่ดีรวมถึง การสร้างวัฒนธรรมในการตรวจประเมินให้เป็นมาตรฐานเดียวกัน โดยพัฒนาการประเมินให้มีความตรง น่าเชื่อถือได้รับการยอมรับมากขึ้น และสอดคล้องกับวัตถุประสงค์การประเมิน การตรวจสอบยืนยันความถูกต้องของงานประเมิน และความโปร่งใสของกลุ่มผู้ประเมิน สามารถให้ข้อมูลที่ถูกต้องย้อนกลับไปยังสถาบันที่มีการตรวจประเมิน

สรุปการจัดฝึกอบรมนักประเมินอภิมานทั้ง 2 รุ่น โดยรุ่นที่ 1 จัดขึ้น ณ อาคารประชุมสุข อาชีวอารุง คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ระหว่างวันที่ 19-20 กันยายน 2552 มีผู้เข้าร่วมอบรม 12 คน และรุ่นที่ 2 จัดขึ้น ณ ห้องบงกช อาคาร 10 ชั้น 6 มหาวิทยาลัยศรีปทุม ระหว่างวันที่ 4-5 มกราคม 2553 มีผู้เข้าร่วมอบรม 16 คน สำหรับการจัดการฝึกอบรม มีส่วนที่เหมือนกัน และส่วนที่แตกต่างกัน ดังต่อไปนี้ คือ

ส่วนที่เหมือนกันในการจัดการฝึกอบรมทั้ง 2 รุ่น มี 7 ข้อ ดังต่อไปนี้

1. หลักสูตรในการฝึกอบรมนักประเมินอภิมาน ใช้หลักสูตรเดียวกันทั้ง 2 รุ่น
2. เอกสารประกอบการอบรม 7 ชุด ได้แก่ 1) คู่มือการประเมินอภิมาน 2) แบบประเมินอภิมาน 3) เอกสารแนวคิดและหลักการประเมินอภิมาน 4) เอกสารมาตรฐานการประเมินอภิมาน และวิธีการขั้นตอนการประเมินอภิมาน 5) เอกสารเทคนิคการประเมินคุณภาพภายในและการเขียนรายงาน 6) รูปแบบการเขียนรายงานการประเมินอภิมาน และ 7) แบบประเมินโครงการฝึกอบรมนักประเมินอภิมาน

3. รายงานการประเมินตนเอง และรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา 5 ฉบับ ได้แก่ มหาวิทยาลัยของรัฐ 1 มหาวิทยาลัยราชภัฏ 1 มหาวิทยาลัยเทคโนโลยีราชมงคล 1 มหาวิทยาลัยเอกชน 1 และมหาวิทยาลัยเอกชน 2

4. วิทยากรหลัก โดยผู้วิจัย และวิทยากรรับเชิญ คือ ดร.วราภรณ์ ไทโยมา

5. ขั้นตอน และกิจกรรมการฝึกอบรม ได้แก่ 1) การบรรยายแนวคิดและหลักการประเมินอภิมาน มาตรฐานและเกณฑ์ในการประเมินอภิมาน และขั้นตอน วิธีการประเมินอภิมาน รวมถึงเทคนิคการประเมินคุณภาพภายในของสถาบันอุดมศึกษา ระดับสถาบันและการเขียนรายงานผลการประเมิน 2) การศึกษามาตรฐาน ตัวบ่งชี้การประเมินอภิมาน 3) ศึกษารายงานการประเมินตนเอง ระดับสถาบัน และรายงานผลการประเมินคุณภาพการศึกษาภายในของสถานศึกษา 4) ฝึกปฏิบัติประเมินอภิมานตามมาตรฐานและเกณฑ์การประเมินอภิมาน และ 5) การเขียนรายงานการประเมินอภิมานและนำเสนอผลการประเมินอภิมาน

6. ผู้เข้าอบรมมีความเข้าใจเกี่ยวกับการประเมินอภิมาน มาตรฐานการประเมินอภิมาน รวมถึงมีทักษะในการประเมินอภิมาน โดยสามารถทำการประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน ได้ตามที่หลักสูตรกำหนด เนื่องจากจุดเน้นของหลักสูตรต้องการให้ผู้เข้าอบรมสามารถทำหน้าที่เป็นนักประเมินอภิมานได้ จึงเน้นทักษะการประเมินและเขียนรายงานเป็นหลัก

7. ระยะเวลาในการจัดการอบรมรวม 2 วัน

ส่วนที่แตกต่างกันในการจัดการฝึกอบรมทั้ง 2 รุ่นหรือส่วนที่มีการปรับปรุงจากการอบรมในรุ่นที่ 1 ดังมีรายละเอียดต่อไปนี้

1. ผู้เข้ารับการอบรมในหลักสูตรฝึกอบรมนักประเมินอภิมาน โดยรุ่นที่ 1 เป็นผู้ที่เกี่ยวข้อง และมีความรู้เกี่ยวกับการประกันคุณภาพภายในของสถาบันอุดมศึกษา หรือเป็นผู้ประเมินคุณภาพภายใน จำนวน 12 คน ส่วนผู้เข้ารับการอบรมในรุ่นที่ 2 เป็นผู้ประเมินคุณภาพภายในของมหาวิทยาลัยศรีปทุม จำนวน 15 คน และมหาวิทยาลัยกรุงเทพ 1 คน

2. มีเอกสารเพิ่มเติมสำหรับผู้เข้ารับการอบรมรุ่น 2 คือ คู่มือการประกันคุณภาพการศึกษาภายในสถานศึกษา ระดับอุดมศึกษา และแจกให้ผู้เข้ารับการอบรมรุ่นที่ 1 ภายหลังจากการอบรมในรุ่นที่ 1 พบว่า การประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน จำเป็นต้องมีคู่มือการประกันคุณภาพการศึกษาภายใน เพื่อใช้ประกอบการตรวจสอบรายงานการประเมินตนเองของสถาบันอุดมศึกษา

3. ปรับปรุงเอกสารที่ใช้การฝึกอบรมรุ่นที่ 2 จากข้อเสนอแนะของผู้เข้ารับการอบรมรุ่นที่ 1 ได้แก่ 1) คู่มือการประเมินอภิมาน ปรับปรุงมาตรฐาน ตัวบ่งชี้ เกณฑ์การประเมิน

เพิ่มเติมนิยามศัพท์ 2) แบบประเมินอภิमान ปรับปรุงตัวบ่งชี้ และเกณฑ์การประเมิน 3) แบบประเมินโครงการฝึกอบรมนักประเมินอภิमान เพิ่มเติมด้านความรู้ ความเข้าใจเกี่ยวกับการประเมินอภิमान

2.3 ผลการวิเคราะห์ความสอดคล้องของผลการประเมินอภิमानของผู้เข้ารับการอบรม

การวิเคราะห์ความสอดคล้องของผลการประเมินอภิमान พิจารณาใน 3 ประเด็น คือ 1) ค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างกลุ่มผู้เข้ารับการอบรม นักประเมินอภิमानทั้ง 2 รุ่น 2) ค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างมหาวิทยาลัยของผู้เข้ารับการอบรมรุ่นที่ 1 3) ค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างมหาวิทยาลัยของผู้เข้ารับการอบรมรุ่นที่ 2 ดังมีรายละเอียดต่อไปนี้

1) ผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างผู้เข้ารับการฝึกอบรมเชิงปฏิบัติการนักประเมินทั้ง 2 รุ่น โดยทำการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา จำนวน 5 ฉบับๆเดียวกัน มีความสอดคล้องกันอย่างมีนัยสำคัญที่ระดับ .05 ($r=0.335$) ดังตารางที่ 4.15

ตารางที่ 4.15 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างผู้เข้ารับการอบรมนักประเมินอภิमानทั้ง 2 รุ่น

Meta-evaluator	Training 1	Training 2
Training 1	1.000	
Training 2	0.335*	1.000
M	2.88	2.43
SD	0.778	0.662

* $p < .05$

2) ผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างมหาวิทยาลัยทั้ง 5 ฉบับ ที่ประเมินโดยผู้เข้ารับการอบรมในรุ่นที่ 1 พบว่า มีความสอดคล้องกันอย่างมีนัยสำคัญที่ระดับ .01 จำนวน 4 คู่ และระดับ .05 จำนวน 2 คู่ โดยค่าสัมประสิทธิ์สหสัมพันธ์ที่มีความสอดคล้องกันสูงสุด คือ ผลการประเมินของมหาวิทยาลัยเอกชน 1 กับมหาวิทยาลัยราชภัฏ 1 มีค่าเท่ากับ 0.657 รองลงมาคือ ผลการประเมินของมหาวิทยาลัยเอกชน 1 กับมหาวิทยาลัยเทคโนโลยีราชมงคล 1 และมหาวิทยาลัยราชภัฏ 1 กับมหาวิทยาลัยเทคโนโลยีราชมงคล 1 มีค่าเท่ากับ 0.590 และ 0.443 ตามลำดับ นอกจากนี้ค่าสัมประสิทธิ์สหสัมพันธ์ที่มีความสอดคล้องกันต่ำสุด คือ ผลการประเมินของมหาวิทยาลัยของรัฐ 1 กับมหาวิทยาลัยเอกชน 2 มีค่าเท่ากับ 0.136 ดังตารางที่ 4.16

ตารางที่ 4.16 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างมหาวิทยาลัยของผู้เข้ารับการอบรมรุ่นที่ 1

Meta-evaluation	ม.ของรัฐ 1	ม.เอกชน 1	ม.เอกชน 2	ม.ราชภัฏ 1	ม.ราชมงคล 1
ม.ของรัฐ 1	1.000				
ม.เอกชน 1	0.345*	1.000			
ม.เอกชน 2	0.136	0.425**	1.000		
ม.ราชภัฏ 1	0.314	0.657**	0.222	1.000	
ม.ราชมงคล 1	0.350*	0.590**	0.257	0.443**	1.000
M	2.95	3.00	2.79	3.03	2.76
SD	1.012	1.162	1.094	1.026	1.261

* $p < .05$, ** $p < .01$

3) ผลการวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างมหาวิทยาลัยทั้ง 5 ฉบับ ที่ประเมินโดยผู้เข้ารับการอบรมในรุ่นที่ 2 พบว่า มีความสอดคล้องกันอย่างมีนัยสำคัญที่ระดับ .01 จำนวน 6 คู่ และระดับ .05 จำนวน 1 คู่ โดยค่าสัมประสิทธิ์สหสัมพันธ์ที่มีความสอดคล้องกันสูงสุด คือ ผลการประเมินของมหาวิทยาลัยของรัฐ 1 กับมหาวิทยาลัยเอกชน 2 มีค่าเท่ากับ 0.595 รองลงมาคือ ผลการประเมินของมหาวิทยาลัยเอกชน 1 กับมหาวิทยาลัยเทคโนโลยีราชมงคล 1 และมหาวิทยาลัยเอกชน 2 กับมหาวิทยาลัยเทคโนโลยีราชมงคล 1 มีค่าเท่ากับ 0.534 และ 0.458 ตามลำดับ นอกจากนี้ค่าสัมประสิทธิ์สหสัมพันธ์ที่มีความสอดคล้องกันต่ำสุด คือ ผลการประเมินของมหาวิทยาลัยเอกชน 1 กับมหาวิทยาลัยราชภัฏ 1 มีค่าเท่ากับ 0.128 ดังตารางที่ 4.17

ตารางที่ 4.17 ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐานและค่าสัมประสิทธิ์สหสัมพันธ์ของผลการประเมินอภิमानระหว่างมหาวิทยาลัยของผู้เข้ารับการอบรมรุ่นที่ 2

Meta-evaluation	ม.ของรัฐ 1	ม.เอกชน 1	ม.เอกชน 2	ม.ราชภัฏ 1	ม.ราชมงคล 1
ม.ของรัฐ 1	1.000				
ม.เอกชน 1	0.436**	1.000			
ม.เอกชน 2	0.595**	0.290	1.000		
ม.ราชภัฏ 1	0.354*	0.128	0.211	1.000	
ม.ราชมงคล 1	0.434**	0.534**	0.458**	0.446**	1.000
M	2.37	2.79	2.47	2.32	2.45
SD	0.998	0.991	1.109	0.962	0.795

* $p < .05$, ** $p < .01$

ตอนที่ 3 ผลการตรวจสอบคุณภาพของเครื่องมือประเมินอภิमान

ผลการตรวจสอบคุณภาพเครื่องมือประเมินอภิमान แบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 ผลการตรวจสอบคุณภาพเครื่องมือด้านความตรง ส่วนที่ 2 ผลการตรวจสอบคุณภาพเครื่องมือด้านความเที่ยง ดังมีรายละเอียดดังต่อไปนี้

ส่วนที่ 1 ผลการตรวจสอบคุณภาพเครื่องมือ ด้านความตรงเชิงโครงสร้าง (construct validity) ด้วยเทคนิคกลุ่มรู้ชัด (known group technique) โดยใช้ผู้ทรงคุณวุฒิด้านการประกันคุณภาพการศึกษา จำนวน 10 ท่าน เป็นผู้บ่งชี้คุณภาพของรายงานผลการประเมินคุณภาพภายใน จำนวน 10 แห่ง โดยสามารถจำแนกรายงานผลการประเมินออกเป็น 2 กลุ่ม กลุ่มสูง 6 ฉบับ และกลุ่มต่ำจำนวน 4 ฉบับ คือ กลุ่มแรก เป็นรายงานที่ผลการประเมินคุณภาพสูงในระดับดีขึ้น และกลุ่มที่สอง เป็นรายงานที่มีผลการประเมินคุณภาพต่ำกว่าระดับดี เพื่อเปรียบเทียบคุณภาพกับผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ที่ประเมินโดยนักประเมินอภิमानที่ผ่านการฝึกอบรมตามหลักสูตรที่พัฒนาขึ้น พบว่า ผลการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ระหว่างนักประเมินอภิमानประเมินรายงานตามกลุ่มที่แบ่งโดยผู้ทรงคุณวุฒิบ่งชี้ กลุ่มคุณภาพทั้งสองกลุ่ม แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ($t = 2.621, p\text{-value} = .031$) ดังตารางที่ 4.18

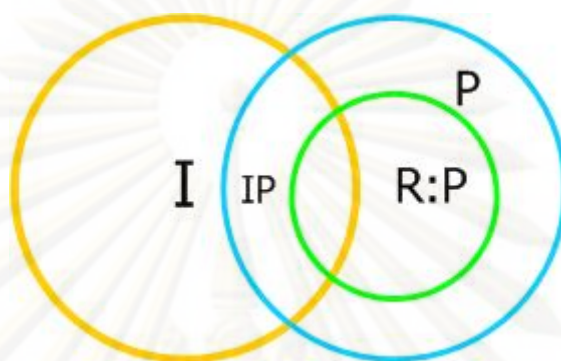
ตารางที่ 4.18 ผลการวิเคราะห์ความแตกต่างของค่าเฉลี่ยผลการประเมินคุณภาพรายงาน ระหว่างกลุ่มนักประเมินอภิमान

กลุ่มที่ประเมิน	ค่าเฉลี่ย			Levene's Test		t-test		
	กลุ่ม	M	SD	F	p-value	t	df	p-value
ผู้ทรงคุณวุฒิ	6	3.017	.45019	.281	.610	3.536	8	.008
	4	2.113	.28395					
นักประเมินอภิमान	6	2.987	.31392	.701	.427	2.621	8	.031
	4	2.503	.23272					

ส่วนที่ 2 ผลการตรวจสอบคุณภาพเครื่องมือด้านความเที่ยง (Reliability) ของแบบประเมินอภิमान จากผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน 3, 5, 7, และ 9 ฉบับ ของนักประเมินอภิमान 1, 2, และ 3 คนต่างกัน

ผลการประมาณค่าความแปรปรวนของแบบประเมินอภิमानที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) จากการประเมินของนักประเมินอภิमान ตั้งแต่ 1, 2, และ 3 คนต่างกัน ที่ทำการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน

จำนวน 3, 5, 7, และ 9 ฉบับ โดยวิธีการหาความเที่ยงโดยวิธีสัมประสิทธิ์การสรุปอ้างอิง (G-Theory) I X (R:P) Design เมื่อทำการประเมินรายงานผลการประเมินคุณภาพภายใน จำนวน 3, 5, 7, และ 9 ฉบับ โดยใช้นักประเมินอภิมานจำนวน 1, 2, และ 3 คนต่างกัน ด้วยแบบประเมินอภิมานเหมือนกัน จำนวน 38 ตัวบ่งชี้ ดังภาพที่ 4.8



ภาพที่ 4.8 ความสัมพันธ์ของวิธีสัมประสิทธิ์การสรุปอ้างอิง แบบ I x (R:P) Design

การประมาณความแปรปรวนของคะแนนการประเมินอภิมานที่มีต่อค่าสัมประสิทธิ์การสรุปอ้างอิง จากการประเมินคุณภาพภายใน พบว่า องค์ประกอบความแปรปรวนของตัวบ่งชี้ในแบบประเมินอภิมาน ($\hat{\sigma}_{(i)}^2$) มีค่ามากที่สุด คิดเป็น 37.43% ของความแปรปรวนทั้งหมด รองลงมา คือ องค์ประกอบความแปรปรวนของจำนวนรายงานการประเมินคุณภาพภายใน ($\hat{\sigma}_{(p)}^2$) และจำนวนนักประเมินอภิมานที่ประเมินรายงานผลการประเมินคุณภาพภายในแต่ละฉบับ ($\hat{\sigma}_{(rp)}^2$) คิดเป็น 34.95% และ 13.38% ของความแปรปรวนทั้งหมด ตามลำดับ ส่วนองค์ประกอบความแปรปรวนที่มีค่าน้อย คือ ผลร่วมระหว่างจำนวนตัวบ่งชี้แบบประเมินอภิมานกับรายงานผลการประเมินคุณภาพภายใน ($\hat{\sigma}_{(ip)}^2$) และส่วนที่เหลือ ($\hat{\sigma}_{(irp)}^2$) ที่เป็นความแปรปรวนของความคลาดเคลื่อนที่เหลือจากทั้ง 4 แหล่ง คิดเป็น 7.25% และ 6.99% ของความแปรปรวนทั้งหมด ตามลำดับ ดังตารางที่ 4.19

ตารางที่ 4.19 ผลการประมาณค่าความแปรปรวนของคะแนนการประเมินอภิมานที่มีผลต่อ

ค่าสัมประสิทธิ์การสรุปอ้างอิง ด้วยการประเมินรายงานผลการประเมินคุณภาพภายใน

Source of Variation	Sum of Square	df	Mean Square	EMS	% Total Variance
I	148.097	37	4.003	4.006	37.432
P	17.605	4	4.401	3.740	34.947
R:P	7.092	5	1.418	1.432	13.381
IP	114.995	148	0.777	0.776	7.251
IR:P	138.408	185	0.748	0.748	6.989
Total	426.197	379	11.347	10.702	100.000

D-study I x (R:P) Design

เมื่อ	I	แทน	จำนวนตัวบ่งชี้ในแบบประเมินอภิमान
	P	แทน	จำนวนรายงานผลการประเมินคุณภาพภายใน
	R:P	แทน	จำนวนนักประเมินอภิमानที่ประเมินรายงานผลการประเมินคุณภาพภายในแต่ละฉบับ
	IP	แทน	ผลร่วมระหว่างจำนวนตัวบ่งชี้แบบประเมินอภิमानกับรายงานผลการประเมินคุณภาพภายใน
	IR:P	แทน	ส่วนที่เหลือ

การประมาณค่าความแปรปรวนของการประเมินอภิमानที่ทำการประเมินรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา จำนวน 3, 5, 7, และ 9 ฉบับ จากนักประเมินอภิमानจำนวน 1, 2, 3 คน ความแปรปรวนของแบบประเมินประเมินอภิमान มีค่าเท่ากับ 0.323 ทุกกลุ่ม โดยค่าความคลาดเคลื่อนเชิงสัมบูรณ์ กรณีใช้รายงานผลประเมินคุณภาพภายใน 3 ฉบับ มีค่า 0.273, 0.145, และ 0.103 ตามลำดับ กรณีเพิ่มรายงานผลประเมินคุณภาพภายในเท่ากับ 5 ฉบับ มีค่า 0.164, 0.087, และ 0.062 ตามลำดับ กรณีเพิ่มรายงานผลประเมินคุณภาพภายในเท่ากับ 7 ฉบับ มีค่า 0.117, 0.062, และ 0.044 ตามลำดับ และกรณีเพิ่มรายงานผลประเมินคุณภาพภายในเป็น 9 ฉบับ มีค่า 0.091, 0.048, และ 0.034 ตามลำดับ แสดงให้เห็นว่า เมื่อเพิ่มจำนวนนักประเมินอภิमान และเพิ่มรายงานผลการประเมินคุณภาพภายในมากขึ้น ความคลาดเคลื่อนจะลดลงตามลำดับ

สัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์ (phi coefficient) ในกรณีใช้ รายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา จำนวน 3, 5, 7, และ 9 ฉบับ และจำนวนนักประเมินอภิमानจำนวน 1, 2, 3 คน กรณีใช้รายงานผลประเมินคุณภาพภายใน 3 ฉบับ มีค่า 0.542, 0.689, และ 0.758 ตามลำดับ หากเพิ่มรายงานผลประเมินคุณภาพภายในเท่ากับ 5 ฉบับ มีค่า 0.663, 0.787, และ 0.839 ตามลำดับ กรณีเพิ่มรายงานผลประเมินคุณภาพภายในเท่ากับ 7 ฉบับ มีค่า 0.734, 0.838, และ 0.880 ตามลำดับ และกรณีเพิ่มรายงานผลประเมินคุณภาพภายในเป็น 9 ฉบับ มีค่า 0.780, 0.869, และ 0.904 ตามลำดับ แสดงให้เห็นว่ากรณี เพิ่มจำนวนรายงานผลการประเมินคุณภาพภายใน และเพิ่มจำนวนนักประเมินอภิमान ค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) มีค่ามากขึ้น จากผลการวิเคราะห์พบว่า ควรกำหนดนักประเมินอภิमानจำนวน 3 คนขึ้นไป กรณีประเมินรายงานผลการประเมินฯ จำนวน 3 ฉบับ และควรกำหนดนักประเมินอภิमानจำนวน 2 คนขึ้นไป สำหรับประเมินรายงานผลการประเมินฯ จำนวน 5 ฉบับ

และนอกจากนี้ควรใช้นักประเมินอภิมานตั้งแต่ 1 คนขึ้นไป สำหรับประเมินรายงานฯ จำนวน 7 และ 9 ฉบับ ตามลำดับ เนื่องจากมีค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์ มากกว่า 0.70 ซึ่งจะ ทำให้การประเมินอภิมานมีประสิทธิภาพ น่าเชื่อถือ และสามารถสรุปอ้างอิงได้ รายละเอียดดัง ตารางที่ 4.20

ตารางที่ 4.20 ผลการประมาณค่าความแปรปรวนของคะแนนการประเมินอภิมาน ความคลาดเคลื่อน และค่าสัมประสิทธิ์การสรุปอ้างอิงในการประเมินรายงานผลการประเมินคุณภาพ ภายใน 3, 5, 7, 9 ฉบับ ด้วยนักประเมินอภิมานระหว่าง 1, 2, 3 คน

EFFECT I x (R:P) Design	ESTIMATED VARIANCE COMPONENT IN D-STUDY												
	n'_p	3			5			7			9		
	n'_r	1	2	3	1	2	3	1	2	3	1	2	3
$\hat{\sigma}_{(i)}^2 = 0.323$	$\hat{\sigma}_{(I)}^2$	0.323	0.323	0.323	0.323	0.323	0.323	0.323	0.323	0.323	0.323	0.323	0.323
$\hat{\sigma}_{(p)}^2 = 0.039$	$\hat{\sigma}_{(P)}^2$	0.013	0.013	0.013	0.008	0.008	0.008	0.006	0.006	0.006	0.004	0.004	0.004
$\hat{\sigma}_{(r:p)}^2 = 0.018$	$\hat{\sigma}_{(R:P)}^2$	0.006	0.003	0.002	0.004	0.002	0.001	0.003	0.001	0.001	0.002	0.001	0.001
$\hat{\sigma}_{(ip)}^2 = 0.014$	$\hat{\sigma}_{(IP)}^2$	0.005	0.005	0.005	0.003	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002
$\hat{\sigma}_{(ir:p)}^2 = 0.748$	$\hat{\sigma}_{(IR:P)}^2$	0.249	0.125	0.083	0.150	0.075	0.050	0.107	0.053	0.036	0.083	0.042	0.028
ERROR	$\hat{\sigma}_{\delta}^2$	0.254	0.130	0.088	0.153	0.078	0.053	0.109	0.056	0.038	0.085	0.043	0.029
VARIANCE	$\hat{\sigma}_{\Delta}^2$	0.273	0.145	0.103	0.164	0.087	0.062	0.117	0.062	0.044	0.091	0.048	0.034
G - COEFFICIENT	$\hat{\rho}_{\delta}^2$	0.559	0.714	0.786	0.679	0.806	0.859	0.748	0.853	0.895	0.792	0.882	0.917
	$\hat{\rho}_{\Delta}^2$	0.542	0.689	0.758	0.663	0.787	0.839	0.734	0.838	0.880	0.780	0.869	0.904

ตอนที่ 4 ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา

ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ผู้วิจัยขอนำเสนอแบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา และส่วนที่ 2 ผลการตรวจสอบและสังเคราะห์เนื้อหาารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ดังมีรายละเอียดดังต่อไปนี้

ส่วนที่ 1 ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา

ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา พิจารณาจากเกณฑ์การตัดสิน ดังต่อไปนี้ คือ ค่าเฉลี่ย 3.26 ขึ้นไป หมายถึง งานประเมินมีคุณภาพดีมาก ค่าเฉลี่ย 2.51 – 3.25 หมายถึง งานประเมินมีคุณภาพดี ค่าเฉลี่ย 1.76 – 2.50 หมายถึง งานประเมินมีคุณภาพพอใช้ และค่าเฉลี่ยน้อยกว่า 1.76 หมายถึง ควรปรับปรุงคุณภาพงานประเมิน

ดังนั้น ผู้วิจัยขอนำเสนอผลการประเมินอภิमानในภาพรวม และจำแนกตามกลุ่มสถาบัน 4 กลุ่ม ได้แก่ มหาวิทยาลัยของรัฐ มหาวิทยาลัยเอกชน มหาวิทยาลัยราชภัฏ และมหาวิทยาลัยเทคโนโลยีราชมงคล แบ่งออกเป็น 3 ประเด็น คือ 1) ผลการประเมินอภิमानของสถาบันอุดมศึกษาในภาพรวม 2) ผลการประเมินอภิमानของสถาบันอุดมศึกษา จำแนกตามกลุ่มสถาบัน และ 3) ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิमान ดังมีรายละเอียดดังต่อไปนี้

1. ผลการประเมินอภิमानของสถาบันอุดมศึกษาในภาพรวม

ผลการประเมินอภิमानของสถาบันอุดมศึกษาในภาพรวม พบว่า มีผลการประเมินอยู่ในระดับดี (M= 2.78, SD=.390) ผลการประเมินภาพรวมสูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ เมื่อพิจารณาผลประเมินอภิमानในรายมาตรฐาน พบว่า มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินสูงสุด อยู่ในระดับดี (M= 3.21, SD=.502) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ รองลงมาคือ มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดี (M= 3.11, SD=.558) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับควรปรับปรุง มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดี (M= 3.08, SD=.527) ผลการประเมินใน

มาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ และมาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี ($M= 2.73, SD=.573$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับควรปรับปรุง ส่วนมาตรฐานด้านประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ ($M= 1.79, SD=.526$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดี และต่ำสุดอยู่ในระดับควรปรับปรุง ดังตารางที่ 4.21

ตารางที่ 4.21 ผลการประเมินอภิमानสถาบันอุดมศึกษาภาพรวม

มาตรฐานการประเมิน	N	Min	Max	M	SD	แปลผล
ความตรงของการประเมิน (Validity)	50	1.40	3.80	2.7327	.57287	ดี
อรรถประโยชน์ (Utility)	50	1.80	3.80	3.0760	.52705	ดี
จริยธรรมในการประเมิน (Ethicality)	50	1.50	3.79	3.1092	.55838	ดี
ความน่าเชื่อถือของนักประเมิน (Credibility)	50	2.25	4.00	3.2050	.50176	ดี
ประสิทธิผลต้นทุน (Cost-effectiveness)	50	1.00	3.00	1.7900	.52576	พอใช้
รวม	50	1.96	3.65	2.7826	.38963	ดี

ผลการประเมินอภิमानของสถาบันอุดมศึกษา จำแนกตามมาตรฐาน พบว่า มาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี ($M= 2.73, SD=.573$) โดยตัวบ่งชี้ Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ มีผลการประเมินสูงสุด ($M= 3.54, SD=.579$) อยู่ในระดับดีมาก รองลงมา คือ Va1.4 การออกแบบการประเมิน ($M= 2.94, SD=1.007$) อยู่ในระดับดี ส่วนตัวบ่งชี้ Va1.8 การตรวจสอบคุณภาพของสารสนเทศ มีผลการประเมินต่ำสุด ($M= 2.12, SD=.689$) อยู่ในระดับพอใช้ มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดี ($M= 3.08, SD=.527$) โดยตัวบ่งชี้ Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม ($M= 3.54, SD=.676$) อยู่ในระดับดีมาก รองลงมา คือ Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา ($M= 3.34, SD=.848$) อยู่ในระดับดีมาก ส่วนตัวบ่งชี้ Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล มีผลการประเมินต่ำสุด ($M= 2.42, SD=.575$) อยู่ในระดับพอใช้ มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดี ($M= 3.11, SD=.558$) โดยตัวบ่งชี้ Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง มีผลการประเมินสูงสุด ($M= 3.56, SD=.812$) อยู่ในระดับดีมาก รองลงมา คือ Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน ($M= 3.40, SD=.639$) อยู่ในระดับดีมากเช่นกัน ส่วนตัวบ่งชี้ Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย มีผลการประเมินต่ำสุด ($M= 2.52, SD=.789$) อยู่ในระดับดี มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินอยู่ในระดับดี ($M= 3.21,$

SD=.502) โดยตัวบ่งชี้ Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน มีผลการประเมินสูงสุด (M= 3.32, SD=.713) อยู่ในระดับดีมาก สำหรับตัวบ่งชี้ที่มีผลการประเมินต่ำสุด คือ Cr4.1 ผู้ประเมินมีความสามารถในการประเมิน และCr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน (M= 3.14, SD=.700, .833) อยู่ในระดับดีเช่นกัน ส่วน**มาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้** (M=1.79, SD=.526) โดยตัวบ่งชี้ Ce5.1 การจัดการทรัพยากร การประเมินอย่างคุ้มค่า (M= 2.24, SD=.771) อยู่ในระดับพอใช้ ซึ่งสูงกว่าตัวบ่งชี้ Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน (M=1.34, SD=.557) ที่อยู่ในระดับควรปรับปรุง ดังตารางที่ 4.22

ตารางที่ 4.22 ผลประเมินอภิमानสถาบันอุดมศึกษา จำแนกตามมาตรฐาน

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Va1.1 การระบุบริบทของสถาบัน	50	1.00	4.00	2.6000	.80812	ดี
Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน	50	1.00	4.00	2.2800	.99057	พอใช้
Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมิน	50	1.00	4.00	2.3400	1.27151	พอใช้
Va1.4 การออกแบบการประเมิน	50	1.00	4.00	2.9200	1.00691	ดี
Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสาร	50	1.00	4.00	2.7400	.80331	ดี
Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	50	1.00	4.00	2.8400	.79179	ดี
Va1.7 การตรวจสอบความถูกต้องของข้อมูล	50	1.00	4.00	2.7600	.84660	ดี
Va1.8 การตรวจสอบคุณภาพของสารสนเทศ	50	1.00	4.00	2.1200	.68928	พอใช้
Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูล	50	1.00	4.00	2.9000	.99488	ดี
Va1.10 การแปลผลและสรุปผลการประเมินถูกต้อง	50	1.00	4.00	2.7400	.87622	ดี
Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ	50	2.00	4.00	3.5400	.57888	ดี
Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง	50	1.00	4.00	2.7800	.95383	ดี
Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมิน	50	1.00	4.00	2.8400	1.14927	ดี
รวมมาตรฐานความตรงของการประเมิน	50	1.40	3.80	2.7327	.57287	ดี
Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียในการประเมิน	50	1.00	4.00	3.0200	1.03982	ดี
Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล	50	1.00	4.00	3.1800	.96235	ดี
Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล	50	1.00	4.00	2.4200	.57463	พอใช้
Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมิน	50	2.00	4.00	3.1400	.67036	ดี
Ut 2.5 ผลการประเมินเป็นประโยชน์ ในการพัฒนาการดำเนินงาน	50	1.00	4.00	3.2200	.78999	ดี
Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด	50	1.00	4.00	2.8600	1.14304	ดี
Ut 2.7 รายงานผลการประเมินครบถ้วน และข้อเสนอแนะเพื่อการพัฒนา	50	1.00	4.00	3.1600	.93372	ดี
Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ	50	1.00	4.00	2.8800	.71827	ดี

ตารางที่ 4.22 (ต่อ)

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา	50	1.00	4.00	3.3400	.84781	ดีมาก
Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม	50	2.00	4.00	3.5400	.67643	ดีมาก
รวมมาตรฐานอรรถประโยชน์	50	1.80	3.80	3.0760	.52705	ดี
Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน	50	2.00	4.00	3.4000	.63888	ดีมาก
Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง	50	1.00	4.00	3.5600	.81215	ดีมาก
Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง	50	1.00	4.00	2.8400	.71027	ดี
Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน	50	1.00	4.00	2.9000	.81441	ดี
Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัด	50	1.00	4.00	3.0400	.87970	ดี
Et 3.6 การกำหนดวิธีการประเมินโดยไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล	50	1.00	4.00	3.3800	.77959	ดีมาก
Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย	50	1.00	4.00	2.5200	.78870	ดี
Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย	50	1.00	4.00	3.3000	.81441	ดีมาก
Et 3.9 การประเมินที่ครบถ้วน ชัดเจน	50	1.00	4.00	3.2200	.91003	ดี
รวมมาตรฐานจริยธรรมในการประเมิน	50	1.50	3.79	3.1092	.55838	ดี
Cr4.1 ผู้ประเมินมีความสามารถในการประเมิน	50	2.00	4.00	3.1400	.70015	ดี
Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน	50	1.00	4.00	3.3200	.71257	ดีมาก
Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมิน	50	2.00	4.00	3.2200	.70826	ดี
Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน	50	1.00	4.00	3.1400	.83324	ดี
รวมมาตรฐานความน่าเชื่อถือของนักประเมิน	50	2.25	4.00	3.2050	.50176	ดี
Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า	50	1.00	4.00	2.2400	.77090	พอใช้
Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน	50	1.00	3.00	1.3400	.55733	ปรับปรุง
รวมมาตรฐานประสิทธิผลของการประเมิน	50	1.00	3.00	1.7900	.52576	พอใช้

2. ผลการประเมินอภิมานของสถาบันอุดมศึกษา จำแนกตามกลุ่มสถาบัน

ผลการประเมินอภิมานของสถาบันอุดมศึกษา จำแนกตามกลุ่มสถาบัน แบ่งออกเป็น 4 กลุ่ม คือ 2.1) ผลการประเมินอภิมานมหาวิทยาลัยของรัฐ 2.2) ผลการประเมินอภิมานมหาวิทยาลัยเอกชน 2.3) ผลการประเมินอภิมานมหาวิทยาลัยราชภัฏ และ 2.4) ผลการประเมินอภิมานมหาวิทยาลัยเทคโนโลยีราชมงคล ดังมีรายละเอียดต่อไปนี้

2.1 ผลการประเมินอภิมานมหาวิทยาลัยของรัฐ

ผลการประเมินอภิมานมหาวิทยาลัยของรัฐ พบว่า มีผลการประเมินอยู่ในระดับดี (M= 2.80, SD=.410) ผลการประเมินมหาวิทยาลัยของรัฐสูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ เมื่อพิจารณาผลประเมินอภิมานในรายมาตรฐาน พบว่า มาตรฐานความ

น่าเชื่อถือของนักประเมินมีผลการประเมินสูงสุด อยู่ในระดับดีมาก ($M= 3.34$, $SD=.550$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ รองลงมาคือ มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดีมากเช่นกัน ($M= 3.26$, $SD=.441$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับดี มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดี ($M= 2.84$, $SD=.833$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับควรปรับปรุง และมาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี ($M= 2.82$, $SD=.420$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับควรปรับปรุง ($M= 1.75$, $SD=.707$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดี และต่ำสุดอยู่ในระดับควรปรับปรุง ดังตารางที่ 4.23

ตารางที่ 4.23 ผลการประเมินอภิमानมหาวิทยาลัยของรัฐ

มาตรฐานการประเมิน	N	Min	Max	M	SD	แปลผล
ความตรงของการประเมิน (Validity)	8	2.20	3.37	2.8208	.42048	ดี
อรรถประโยชน์ (Utility)	8	2.60	3.80	3.2625	.44058	ดีมาก
จริยธรรมในการประเมิน (Ethicality)	8	1.50	3.58	2.8385	.83317	ดี
ความน่าเชื่อถือของนักประเมิน (Credibility)	8	2.25	4.00	3.3438	.54996	ดีมาก
ประสิทธิผลต้นทุน (Cost-effectiveness)	8	1.00	3.00	1.7500	.70711	ปรับปรุง
รวม	8	2.04	3.30	2.8031	.40977	ดี

ผลการประเมินอภิमानมหาวิทยาลัยของรัฐ จำแนกตามมาตรฐาน พบว่า มาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี ($M= 2.82$, $SD=.420$) โดยตัวบ่งชี้ Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ มีผลการประเมินสูงสุด ($M= 3.75$, $SD=.463$) อยู่ในระดับดีมาก รองลงมา คือ Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง ($M= 3.63$, $SD=.518$) อยู่ในระดับดีมาก ส่วนตัวบ่งชี้ Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน มีผลการประเมินต่ำสุด ($M= 1.63$, $SD=.518$) อยู่ในระดับควรปรับปรุง มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดีมาก ($M= 3.26$, $SD=.441$) โดยตัวบ่งชี้ Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม ($M= 3.88$, $SD=.354$) อยู่ในระดับดีมาก รองลงมา คือ Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา ($M= 3.75$, $SD=.463$) อยู่ในระดับดีมาก ส่วนตัวบ่งชี้ Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล มีผลการประเมินต่ำสุด ($M=$

2.50, SD=.756) อยู่ในระดับพอใช้ **มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดี** (M= 2.84, SD=.833) โดยตัวบ่งชี้ Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง มีผลการประเมินสูงสุด (M= 3.25, SD=1.165) อยู่ในระดับดี รองลงมา คือ Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน, Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัด, และ Et 3.9 การประเมินที่ครบถ้วน ถูกต้อง ชัดเจน (M= 3.13, SD=.835, .991 และ1.126) อยู่ในระดับดีเช่นกัน ส่วนตัวบ่งชี้ Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย มีผลการประเมินต่ำสุด (M= 2.25, SD=1.135) อยู่ในระดับพอใช้ **มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินอยู่ในระดับดีมาก** (M= 3.34, SD=.550) โดยตัวบ่งชี้ Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน มีผลการประเมินสูงสุด (M= 3.75, SD=.463) อยู่ในระดับดีมาก สำหรับตัวบ่งชี้ที่มีผลการประเมินต่ำสุด คือ Cr4.1 ผู้ประเมินมีความสามารถในการประเมิน และ Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน (M= 3.13, SD=.641, .991) อยู่ในระดับดีเช่นกัน ส่วน**มาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับปรับปรุง** (M= 1.75, SD=.707) โดยตัวบ่งชี้ Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า (M= 2.13, SD=.991) อยู่ในระดับพอใช้ ซึ่งสูงกว่า ตัวบ่งชี้ Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน (M= 1.38, SD=.744) ที่อยู่ในระดับควรปรับปรุง ดังตารางที่ 4.24

ตารางที่ 4.24 ผลประเมินอภิมานมหาวิทยาลัยของรัฐ จำแนกตามมาตรฐาน

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Va1.1 การระบุบริบทของสถาบัน	8	2.00	4.00	2.7500	.70711	ดี
Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน	8	1.00	2.00	1.6250	.51755	ปรับปรุง
Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมิน	8	1.00	4.00	2.8750	1.12599	ดี
Va1.4 การออกแบบการประเมิน	8	1.00	4.00	2.8750	1.12599	ดี
Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสาร	8	2.00	4.00	2.8750	.83452	ดี
Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	8	1.00	4.00	2.7500	1.16496	ดี
Va1.7 การตรวจสอบความถูกต้องของข้อมูล	8	2.00	4.00	3.0000	.75593	ดี
Va1.8 การตรวจสอบคุณภาพของสารสนเทศ	8	2.00	3.00	2.3750	.51755	พอใช้
Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูล	8	1.00	4.00	2.7500	1.03510	ดี
Va1.10 การแปลผลและสรุปผลการประเมินถูกต้อง	8	1.00	4.00	2.5000	1.19523	พอใช้
Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ	8	3.00	4.00	3.7500	.46291	ดีมาก
Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง	8	3.00	4.00	3.6250	.51755	ดีมาก
Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมิน	8	1.00	4.00	3.0000	1.30931	ดี
รวมมาตรฐานความตรงของการประเมิน	8	2.20	3.37	2.8208	.42048	ดี

ตารางที่ 4.24 (ต่อ)

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียในการประเมิน	8	3.00	4.00	3.3750	.51755	ดีมาก
Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล	8	1.00	4.00	3.0000	1.06904	ดี
Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล	8	2.00	4.00	2.5000	.75593	พอใช้
Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมิน	8	2.00	4.00	3.2500	.70711	ดี
Ut 2.5 ผลการประเมินเป็นประโยชน์ในการพัฒนาการดำเนินงาน	8	3.00	4.00	3.6250	.51755	ดีมาก
Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด	8	1.00	4.00	2.8750	1.35620	ดี
Ut 2.7 รายงานผลการประเมินครบถ้วน และข้อเสนอแนะเพื่อการพัฒนา	8	1.00	4.00	3.2500	1.03510	ดี
Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ	8	2.00	4.00	3.1250	.83452	ดี
Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา	8	3.00	4.00	3.7500	.46291	ดีมาก
Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม	8	3.00	4.00	3.8750	.35355	ดีมาก
รวมมาตรฐานอรรถประโยชน์	8	2.60	3.80	3.2625	.44058	ดีมาก
Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน	8	2.00	4.00	3.1250	.83452	ดี
Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง	8	1.00	4.00	3.2500	1.16496	ดี
Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง	8	1.00	4.00	2.3750	1.06066	พอใช้
Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน	8	1.00	3.00	2.3750	.74402	พอใช้
Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัด	8	2.00	4.00	3.1250	.99103	ดี
Et 3.6 การกำหนดวิธีการประเมินโดยไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล	8	2.00	4.00	3.0000	.75593	ดี
Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลายหลาย	8	1.00	4.00	2.2500	1.03510	พอใช้
Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย	8	1.00	4.00	3.0000	1.06904	ดี
Et 3.9 การประเมินที่ครบถ้วน ชัดเจน	8	1.00	4.00	3.1250	1.12599	ดี
รวมมาตรฐานจริยธรรมในการประเมิน	8	1.50	3.58	2.8385	.83317	ดี
Cr 4.1 ผู้ประเมินมีความสามารถ ในการประเมิน	8	2.00	4.00	3.1250	.64087	ดี
Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน	8	3.00	4.00	3.7500	.46291	ดีมาก
Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมิน	8	2.00	4.00	3.3750	.74402	ดีมาก
Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน	8	1.00	4.00	3.1250	.99103	ดี
รวมมาตรฐานความน่าเชื่อถือของนักประเมิน	8	2.25	4.00	3.3438	.54996	ดีมาก
Ce 5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า	8	1.00	4.00	2.1250	.99103	พอใช้
Ce 5.2 การจัดสรรงบประมาณสำหรับการประเมิน	8	1.00	3.00	1.3750	.74402	ปรับปรุง
รวมมาตรฐานประสิทธิผลของการประเมิน	8	1.00	3.00	1.7500	.70711	ปรับปรุง

2.2 ผลการประเมินอภิमानมหาวิทยาลัยเอกชน

ผลการประเมินอภิमानมหาวิทยาลัยเอกชน พบว่า มีผลการประเมินอยู่ในระดับดี ($M= 2.72$, $SD=.409$) ผลการประเมินมหาวิทยาลัยเอกชนรวมสูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ เมื่อพิจารณาผลประเมินอภิमानในรายมาตรฐาน พบว่า **มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินสูงสุด อยู่ในระดับดี** ($M= 3.22$, $SD=.533$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับควรปรับปรุง รองลงมาคือ **มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดี** ($M= 3.11$, $SD=.556$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับควรปรับปรุง **มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดี** ($M= 2.93$, $SD=.551$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ และ**มาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี** ($M= 2.67$, $SD=.559$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับควรปรับปรุง ส่วน**มาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับควรปรับปรุง** ($M= 1.68$, $SD=.478$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดี และต่ำสุดอยู่ในระดับควรปรับปรุง ดังตารางที่ 4.25

ตารางที่ 4.25 ผลการประเมินอภิमानมหาวิทยาลัยเอกชน

มาตรฐานการประเมิน	N	Min	Max	M	SD	แปลผล
ความตรงของการประเมิน (Validity)	19	1.67	3.73	2.6667	.55943	ดี
อรรถประโยชน์ (Utility)	19	2.00	3.80	2.9263	.55060	ดี
จริยธรรมในการประเมิน (Ethicality)	19	1.54	3.79	3.1075	.55578	ดี
ความน่าเชื่อถือของนักประเมิน (Credibility)	19	2.25	4.00	3.2237	.53291	ดี
ประสิทธิผลต้นทุน (Cost-effectiveness)	19	1.00	3.00	1.6842	.47757	ปรับปรุง
รวม	19	2.18	3.65	2.7217	.40903	ดี

ผลการประเมินอภิमानมหาวิทยาลัยเอกชน จำแนกตามมาตรฐาน พบว่า **มาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี** ($M= 2.67$, $SD=.559$) โดยตัวบ่งชี้ Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ มีผลการประเมินสูงสุด ($M= 3.53$, $SD=.513$) อยู่ในระดับดีมาก รองลงมา คือ Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ ($M= 3.05$, $SD=.780$) อยู่ในระดับดี ส่วนตัวบ่งชี้ Va1.8 การตรวจสอบคุณภาพของสารสนเทศ มีผลการประเมินต่ำสุด ($M= 2.05$, $SD=.705$) อยู่ในระดับพอใช้ **มาตรฐาน**

อรรถประโยชน์ มีผลการประเมินอยู่ในระดับดี ($M= 2.93, SD=.551$) โดยตัวบ่งชี้ Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม ($M= 3.42, SD=.692$) อยู่ในระดับดีมาก รองลงมา คือ Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา ($M= 3.16, SD=.898$) อยู่ในระดับดี ส่วนตัวบ่งชี้ Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล, และ Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด มีผลการประเมินต่ำสุด ($M= 2.42, SD=.507, 1.261$) อยู่ในระดับพอใช้ **มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดี** ($M= 3.11, SD=.556$) โดยตัวบ่งชี้ Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง มีผลการประเมินสูงสุด ($M= 3.63, SD=.831$) อยู่ในระดับดีมาก รองลงมา คือ Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย ($M= 3.53, SD=.697$) อยู่ในระดับดีมากเช่นกัน ส่วนตัวบ่งชี้ Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย มีผลการประเมินต่ำสุด ($M= 2.47, SD=.772$) อยู่ในระดับพอใช้ **มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินอยู่ในระดับดี** ($M= 3.22, SD=.533$) โดยตัวบ่งชี้ Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน มีผลการประเมินสูงสุด ($M= 3.26, SD=.806$) อยู่ในระดับดีมาก นอกจากนี้ ตัวบ่งชี้อื่นๆ มีผลประเมินอยู่ในระดับดี ($M= 3.21, SD=.713, .855, .787$) ส่วน**มาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับควรปรับปรุง** ($M= 1.68, SD=.478$) โดยตัวบ่งชี้ Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า ($M= 2.16, SD=.765$) อยู่ในระดับพอใช้ ซึ่งสูงกว่า ตัวบ่งชี้ Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน ($M= 1.21, SD=.419$) ที่อยู่ในระดับควรปรับปรุง ดังตารางที่ 4.26

ตารางที่ 4.26 ผลประเมินอภิมานมหาวิทยาลัยเอกชน จำแนกตามมาตรฐาน

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Va1.1 การระบุบริบทของสถาบัน	19	1.00	4.00	2.4211	.90159	พอใช้
Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน	19	1.00	4.00	2.3684	1.06513	พอใช้
Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมิน	19	1.00	4.00	2.1053	1.19697	พอใช้
Va1.4 การออกแบบการประเมิน	19	1.00	4.00	2.8421	1.06787	ดี
Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสาร	19	1.00	4.00	2.7368	.80568	ดี
Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	19	2.00	4.00	3.0526	.77986	ดี
Va1.7 การตรวจสอบความถูกต้องของข้อมูล	19	1.00	4.00	2.5263	.96427	ดี
Va1.8 การตรวจสอบคุณภาพของสารสนเทศ	19	1.00	4.00	2.0526	.70504	พอใช้
Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูล	19	1.00	4.00	2.9474	1.02598	ดี
Va1.10 การแปลผลและสรุปผลการประเมินถูกต้อง	19	1.00	4.00	2.6842	.94591	ดี

ตารางที่ 4.26 (ต่อ)

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ	19	3.00	4.00	3.5263	.51299	ดีมาก
Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง	19	1.00	4.00	2.6316	.95513	ดี
Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมิน	19	1.00	4.00	2.7895	1.22832	ดี
รวมมาตรฐานความตรงของการประเมิน	19	1.67	3.73	2.6667	.55943	ดี
Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียในการประเมิน	19	1.00	4.00	2.9474	1.17727	ดี
Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล	19	1.00	4.00	2.9474	1.12909	ดี
Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล	19	2.00	3.00	2.4211	.50726	พอใช้
Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมิน	19	2.00	4.00	3.0000	.74536	ดี
Ut 2.5 ผลการประเมินเป็นประโยชน์ ในการพัฒนาการดำเนินงาน	19	2.00	4.00	3.0526	.77986	ดี
Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด	19	1.00	4.00	2.4211	1.26121	พอใช้
Ut 2.7 รายงานผลการประเมินครบถ้วน และข้อเสนอแนะเพื่อการพัฒนา	19	1.00	4.00	2.9474	1.02598	ดี
Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ	19	2.00	4.00	2.9474	.62126	ดี
Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา	19	1.00	4.00	3.1579	.89834	ดี
Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม	19	2.00	4.00	3.4211	.69248	ดีมาก
รวมมาตรฐานอรรถประโยชน์	19	2.00	3.80	2.9263	.55060	ดี
Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน	19	2.00	4.00	3.4211	.60698	ดีมาก
Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง	19	1.00	4.00	3.6316	.83070	ดีมาก
Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง	19	2.00	4.00	2.7368	.56195	ดี
Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน	19	2.00	4.00	2.8421	.76472	ดี
Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัด	19	1.00	4.00	2.7895	.97633	ดี
Et 3.6 การกำหนดวิธีการประเมินโดยไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล	19	1.00	4.00	3.4737	.90483	ดีมาก
Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย	19	1.00	4.00	2.4737	.77233	พอใช้
Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย	19	2.00	4.00	3.5263	.69669	ดีมาก
Et 3.9 การประเมินที่ครบถ้วน ชัดเจน	19	2.00	4.00	3.2632	.87191	ดีมาก
รวมมาตรฐานจริยธรรมในการประเมิน	19	1.54	3.79	3.1075	.55578	ดี
Cr4.1 ผู้ประเมินมีความสามารถ ในการประเมิน	19	2.00	4.00	3.2105	.71328	ดี
Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน	19	1.00	4.00	3.2105	.85498	ดี
Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมิน	19	2.00	4.00	3.2105	.78733	ดี
Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน	19	1.00	4.00	3.2632	.80568	ดีมาก

ตารางที่ 4.26 (ต่อ)

มาตรฐาน	N	Min	Max	M	SD	แปลผล
รวมมาตรฐานความน่าเชื่อถือของนักประเมิน	19	2.25	4.00	3.2237	.53291	ดี
Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า	19	1.00	4.00	2.1579	.76472	พอใช้
Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน	19	1.00	2.00	1.2105	.41885	ปรับปรุง
รวมมาตรฐานประสิทธิผลของการประเมิน	19	1.00	3.00	1.6842	.47757	ปรับปรุง

2.3 ผลการประเมินอภิमानมหาวิทยาลัยราชภัฏ

ผลการประเมินอภิमानมหาวิทยาลัยราชภัฏ พบว่า มีผลการประเมินอยู่ในระดับดี (M= 2.79, SD=.356) ผลการประเมินมหาวิทยาลัยราชภัฏรวมสูงสุดอยู่ในระดับดี และต่ำสุดอยู่ในระดับพอใช้ เมื่อพิจารณาผลประเมินอภิमानในรายมาตรฐาน พบว่า มาตรฐานจริยธรรมในการประเมินมีค่าสูงสุด ผลการประเมินอยู่ในระดับดี (M= 3.18, SD=.460) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ รองลงมาคือ มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินอยู่ในระดับดี (M= 3.11, SD=.471) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดี (M= 3.07, SD=.522) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ และมาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี (M= 2.72, SD=.657) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับควรปรับปรุง ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ (M= 1.86, SD=.479) ผลการประเมินในมาตรฐานนี้อยู่ในระดับดี และต่ำสุดอยู่ในระดับควรปรับปรุง ดังตารางที่ 4.27

ตารางที่ 4.27 ผลการประเมินอภิमानมหาวิทยาลัยราชภัฏ

มาตรฐานการประเมิน	N	Min	Max	M	SD	แปลผล
ความตรงของการประเมิน (Validity)	18	1.40	3.80	2.7204	.65670	ดี
อรรถประโยชน์ (Utility)	18	1.80	3.70	3.0722	.52222	ดี
จริยธรรมในการประเมิน (Ethicality)	18	2.00	3.67	3.1782	.45951	ดี
ความน่าเชื่อถือของนักประเมิน (Credibility)	18	2.25	4.00	3.1111	.47140	ดี
ประสิทธิผลต้นทุน (Cost-effectiveness)	18	1.00	3.00	1.8611	.47914	พอใช้
รวม	18	1.96	3.20	2.7886	.35599	ดี

ผลการประเมินอภิमानมหาวิทยาลัยราชภัฏ จำแนกตามมาตรฐาน พบว่า มาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี ($M = 2.72, SD = .657$) โดยตัวบ่งชี้ Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ มีผลการประเมินสูงสุด ($M = 3.44, SD = .705$) อยู่ในระดับดีมาก รองลงมา คือ Va1.4 การออกแบบการประเมิน ($M = 3.00, SD = .970$) อยู่ในระดับดี ส่วนตัวบ่งชี้ Va1.8 การตรวจสอบคุณภาพของสารสนเทศ มีผลการประเมินต่ำสุด ($M = 1.94, SD = .639$) อยู่ในระดับพอใช้ มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดี ($M = 3.07, SD = .522$) โดยตัวบ่งชี้ Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม ($M = 3.44, SD = .784$) อยู่ในระดับดีมาก รองลงมา คือ Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล ($M = 3.33, SD = .767$) อยู่ในระดับดีมากเช่นกัน ส่วนตัวบ่งชี้ Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล มีผลการประเมินต่ำสุด ($M = 2.33, SD = .594$) อยู่ในระดับพอใช้ มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดี ($M = 3.18, SD = .460$) โดยตัวบ่งชี้ Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง มีผลการประเมินสูงสุด ($M = 3.61, SD = .698$) อยู่ในระดับดีมาก รองลงมา คือ Et 3.6 การกำหนดวิธีการประเมินโดยไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล ($M = 3.56, SD = .616$) อยู่ในระดับดีมากเช่นกัน ส่วนตัวบ่งชี้ Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย มีผลการประเมินต่ำสุด ($M = 2.67, SD = .767$) อยู่ในระดับดี มาตรฐานความน่าเชื่อถือของนักประเมิน มีผลการประเมินอยู่ในระดับดี ($M = 3.11, SD = .471$) โดยตัวบ่งชี้ Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน มีผลการประเมินสูงสุด ($M = 3.28, SD = .575$) อยู่ในระดับดีมาก สำหรับตัวบ่งชี้ที่มีผลการประเมินต่ำสุด คือ Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน ($M = 3.00, SD = .840$) อยู่ในระดับดี ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ ($M = 1.86, SD = .479$) โดยตัวบ่งชี้ Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า ($M = 2.33, SD = .767$) อยู่ในระดับพอใช้ ซึ่งสูงกว่า ตัวบ่งชี้ Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน ($M = 1.39, SD = .502$) ที่อยู่ในระดับควรปรับปรุง ดังตารางที่ 4.28

ตารางที่ 4.28 ผลประเมินอภิमानมหาวิทยาลัยราชภัฏ จำแนกตามมาตรฐาน

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Va1.1 การระบุบริบทของสถาบัน	18	1.00	4.00	2.7222	.82644	ดี
Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน	18	1.00	4.00	2.3333	1.02899	พอใช้
Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมิน	18	1.00	4.00	2.3333	1.41421	พอใช้
Va1.4 การออกแบบการประเมิน	18	1.00	4.00	3.0000	.97014	ดี
Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสาร	18	1.00	4.00	2.7222	.82644	ดี
Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	18	1.00	4.00	2.6667	.68599	ดี
Va1.7 การตรวจสอบความถูกต้องของข้อมูล	18	2.00	4.00	2.8333	.78591	ดี
Va1.8 การตรวจสอบคุณภาพของสารสนเทศ	18	1.00	3.00	1.9444	.63914	พอใช้
Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูล	18	1.00	4.00	2.8889	1.02262	ดี
Va1.10 การแปลผลและสรุปผลการประเมินถูกต้อง	18	2.00	4.00	2.8889	.75840	ดี
Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ	18	2.00	4.00	3.4444	.70479	ดีมาก
Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง	18	1.00	4.00	2.5000	.98518	พอใช้
Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมิน	18	1.00	4.00	2.7222	1.12749	ดี
รวมมาตรฐานความตรงของการประเมิน	18	1.40	3.80	2.7204	.65670	ดี
Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียในการประเมิน	18	1.00	4.00	2.9444	1.05564	ดี
Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล	18	2.00	4.00	3.3333	.76696	ดีมาก
Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล	18	1.00	3.00	2.3333	.59409	พอใช้
Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมิน	18	2.00	4.00	3.1111	.58298	ดี
Ut 2.5 ผลการประเมินเป็นประโยชน์ในการพัฒนาการดำเนินงาน	18	1.00	4.00	3.2222	.87820	ดี
Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด	18	2.00	4.00	3.1111	.90025	ดี
Ut 2.7 รายงานผลการประเมินครบถ้วน และข้อเสนอแนะเพื่อการพัฒนา	18	2.00	4.00	3.2222	.80845	ดี
Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ	18	1.00	4.00	2.7222	.82644	ดี
Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา	18	1.00	4.00	3.2778	.95828	ดีมาก
Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม	18	2.00	4.00	3.4444	.78382	ดีมาก
รวมมาตรฐานอรรถประโยชน์	18	1.80	3.70	3.0722	.52222	ดี
Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน	18	2.00	4.00	3.3889	.60768	ดีมาก
Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง	18	2.00	4.00	3.6111	.69780	ดีมาก
Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง	18	2.00	4.00	3.0000	.59409	ดี
Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน	18	1.00	4.00	3.0556	.87260	ดี

ตารางที่ 4.26 (ต่อ)

มาตรฐาน	N	Min	Max	M	SD	แปลผล
Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัด	18	2.00	4.00	3.2222	.73208	ดี
Et 3.6 การกำหนดวิธีการประเมินโดยไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล	18	2.00	4.00	3.5556	.61570	ดีมาก
Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย	18	1.00	4.00	2.6667	.76696	ดี
Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย	18	2.00	4.00	3.1111	.83235	ดี
Et 3.9 การประเมินที่ครบถ้วน ชัดเจน	18	1.00	4.00	3.1667	.98518	ดี
รวมมาตรฐานจริยธรรมในการประเมิน	18	2.00	3.67	3.1782	.45951	ดี
Cr4.1 ผู้ประเมินมีความสามารถในการประเมิน	18	2.00	4.00	3.0556	.72536	ดี
Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน	18	2.00	4.00	3.2778	.57451	ดีมาก
Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมิน	18	2.00	4.00	3.1111	.67640	ดี
Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน	18	1.00	4.00	3.0000	.84017	ดี
รวมมาตรฐานความน่าเชื่อถือของนักประเมิน	18	2.25	4.00	3.1111	.47140	ดี
Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า	18	1.00	4.00	2.3333	.76696	พอใช้
Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน	18	1.00	2.00	1.3889	.50163	ปรับปรุง
รวมมาตรฐานประสิทธิผลของการประเมิน	18	1.00	3.00	1.8611	.47914	พอใช้

2.4 ผลการประเมินอภิमानมหาวิทยาลัยเทคโนโลยีราชมงคล

ผลการประเมินอภิमानมหาวิทยาลัยเทคโนโลยีราชมงคล พบว่า มีผลการประเมินอยู่ในระดับดี ($M = 2.96$, $SD = .465$) ผลการประเมินมหาวิทยาลัยเทคโนโลยีราชมงคลรวมสูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ เมื่อพิจารณาผลประเมินอภิमानในรายมาตรฐาน พบว่า อรรถประโยชน์ มีผลการประเมินอยู่ในระดับดีมาก ($M = 3.36$, $SD = .513$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับดี มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดีมากเช่นกัน ($M = 3.30$, $SD = .322$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับดี รองลงมาคือ มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินอยู่ในระดับดี ($M = 3.25$, $SD = .500$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ และ มาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดีเช่นกัน ($M = 2.89$, $SD = .635$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดีมาก และต่ำสุดอยู่ในระดับพอใช้ ส่วน มาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ ($M = 2.00$, $SD = .612$) ผลการประเมินในมาตรฐานนี้สูงสุดอยู่ในระดับดี และต่ำสุดอยู่ในระดับควรปรับปรุง ดังตารางที่ 4.29

ตารางที่ 4.29 ผลการประเมินอภิमानมหาวิทยาลัยเทคโนโลยีราชมงคล

มาตรฐานการประเมิน	N	Min	Max	M	SD	แปลผล
ความตรงของการประเมิน (Validity)	5	2.17	3.63	2.8867	.63491	ดี
อรรถประโยชน์ (Utility)	5	2.60	3.80	3.3600	.51284	ดีมาก
จริยธรรมในการประเมิน (Ethicality)	5	2.75	3.54	3.3000	.32194	ดีมาก
ความน่าเชื่อถือของนักประเมิน (Credibility)	5	2.50	3.75	3.2500	.50000	ดี
ประสิทธิผลต้นทุน (Cost-effectiveness)	5	1.50	3.00	2.0000	.61237	พอใช้
รวม	5	2.30	3.54	2.9593	.46547	ดี

ผลการประเมินอภิमानมหาวิทยาลัยเทคโนโลยีราชมงคล จำแนกตามมาตรฐานพบว่า **มาตรฐานความตรงของการประเมินมีผลการประเมินอยู่ในระดับดี** (M= 2.89, SD=.635) โดยตัวบ่งชี้ Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ มีผลการประเมินสูงสุด (M= 3.60, SD=.548) อยู่ในระดับดีมาก รองลงมา คือ Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมิน (M= 3.20, SD=.837) อยู่ในระดับดี ส่วนตัวบ่งชี้ Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมิน มีผลการประเมินต่ำสุด (M= 2.40, SD=1.342) อยู่ในระดับพอใช้ **มาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดีมาก** (M= 3.36, SD=.513) โดยตัวบ่งชี้ Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล และ Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม (M= 3.80, SD=.447) อยู่ในระดับดีมาก รองลงมา คือ Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมิน, Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด, Ut 2.7 รายงานผลการประเมินครบถ้วน และข้อเสนอแนะเพื่อการพัฒนา, และ Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา (M= 3.60, SD=.548) อยู่ในระดับดีมาก ส่วนตัวบ่งชี้ Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล มีผลการประเมินต่ำสุด (M= 2.60, SD=.548) อยู่ในระดับดี **มาตรฐานจริยธรรมในการประเมินมีผลการประเมินอยู่ในระดับดีมาก** (M= 3.30, SD=.322) โดยตัวบ่งชี้ Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน มีผลการประเมินสูงสุด (M= 3.80, SD=.447) อยู่ในระดับดีมาก รองลงมา คือ Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง, และ Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย (M= 3.60, SD=.548) อยู่ในระดับดีมากเช่นกัน ส่วนตัวบ่งชี้ Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่างกันหลากหลาย มีผลการประเมินต่ำสุด (M= 2.60, SD=.548) อยู่ในระดับดี **มาตรฐานความน่าเชื่อถือของนักประเมินมีผลการประเมินอยู่ในระดับดี** (M= 3.25, SD=.500) โดยตัวบ่งชี้

Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมิน มีผลการประเมินสูงสุด (M= 3.40, SD=.548) อยู่ในระดับดีมาก สำหรับตัวบ่งชี้อื่นๆ มีผลการประเมินอยู่ในระดับดีเท่ากัน (M=3.20, SD=.837) ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ (M= 2.00, SD=.612) โดยตัวบ่งชี้ Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า (M= 2.40, SD=.548) อยู่ในระดับพอใช้ ซึ่งสูงกว่า ตัวบ่งชี้ Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน (M= 1.60, SD=.894) ที่อยู่ในระดับควรปรับปรุง ดังตารางที่ 4.30

ตารางที่ 4.30 ผลประเมินอภิมานมหาวิทยาลัยเทคโนโลยีราชมงคล จำแนกตามมาตรฐาน

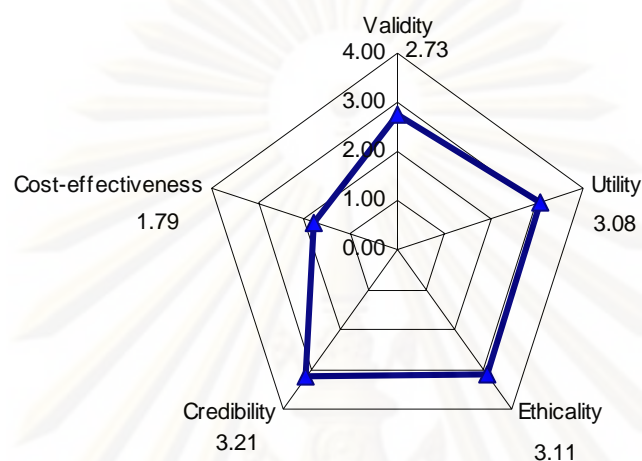
มาตรฐาน	N	Min	Max	M	SD	แปลผล
Va1.1 การระบุบริบทของสถาบัน	5	2.00	3.00	2.6000	.54772	ดี
Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน	5	2.00	4.00	2.8000	.83666	ดี
Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมิน	5	1.00	4.00	2.4000	1.34164	พอใช้
Va1.4 การออกแบบการประเมิน	5	2.00	4.00	3.0000	1.00000	ดี
Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสาร	5	2.00	4.00	2.6000	.89443	ดี
Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	5	2.00	3.00	2.8000	.44721	ดี
Va1.7 การตรวจสอบความถูกต้องของข้อมูล	5	2.00	4.00	3.0000	.70711	ดี
Va1.8 การตรวจสอบคุณภาพของสารสนเทศ	5	2.00	4.00	2.6000	.89443	ดี
Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูล	5	2.00	4.00	3.0000	1.00000	ดี
Va1.10 การแปลผลและสรุปผลการประเมินถูกต้อง	5	2.00	3.00	2.8000	.44721	ดี
Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ	5	3.00	4.00	3.6000	.54772	ดีมาก
Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง	5	2.00	4.00	3.0000	.70711	ดี
Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมิน	5	2.00	4.00	3.2000	.83666	ดี
รวมมาตรฐานความตรงของการประเมิน	5	2.17	3.63	2.8867	.63491	ดี
Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียในการประเมิน	5	1.00	4.00	3.0000	1.22474	ดี
Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูล	5	3.00	4.00	3.8000	.44721	ดีมาก
Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล	5	2.00	3.00	2.6000	.54772	ดี
Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมิน	5	3.00	4.00	3.6000	.54772	ดีมาก
Ut 2.5 ผลการประเมินเป็นประโยชน์ ในการพัฒนาการดำเนินงาน	5	2.00	4.00	3.2000	.83666	ดี
Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด	5	3.00	4.00	3.6000	.54772	ดีมาก
Ut 2.7 รายงานผลการประเมินครบถ้วน และข้อเสนอแนะเพื่อการพัฒนา	5	2.00	4.00	3.6000	.89443	ดีมาก
Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ	5	2.00	3.00	2.8000	.44721	ดี
Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา	5	3.00	4.00	3.6000	.54772	ดีมาก
Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม	5	3.00	4.00	3.8000	.44721	ดีมาก

ตารางที่ 4.30 (ต่อ)

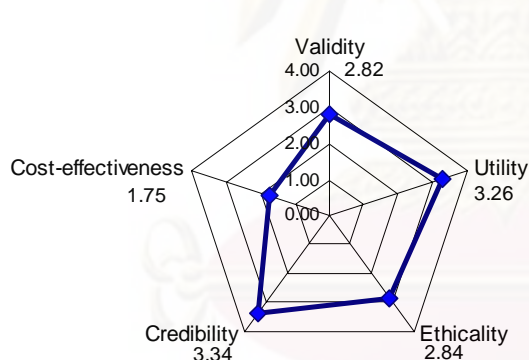
มาตรฐาน	N	Min	Max	M	SD	
รวมมาตรฐานอรรถประโยชน์	5	2.60	3.80	3.3600	.51284	ดีมาก
Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมิน	5	3.00	4.00	3.8000	.44721	ดีมาก
Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง	5	3.00	4.00	3.6000	.54772	ดีมาก
Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง	5	3.00	4.00	3.4000	.54772	ดีมาก
Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน	5	3.00	4.00	3.4000	.54772	ดีมาก
Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัด	5	2.00	4.00	3.2000	.83666	ดี
Et 3.6 การกำหนดวิธีการประเมินโดยไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล	5	2.00	4.00	3.0000	.70711	ดี
Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง ที่แตกต่าง หลากหลาย	5	2.00	3.00	2.6000	.54772	ดี
Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย	5	3.00	4.00	3.6000	.54772	ดีมาก
Et 3.9 การประเมินที่ครบถ้วน ชัดเจน	5	3.00	4.00	3.4000	.54772	ดีมาก
รวมมาตรฐานจริยธรรมในการประเมิน	5	2.75	3.54	3.3000	.32194	ดีมาก
Cr4.1 ผู้ประเมินมีความสามารถในการประเมิน	5	2.00	4.00	3.2000	.83666	ดี
Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน	5	2.00	4.00	3.2000	.83666	ดี
Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมิน	5	3.00	4.00	3.4000	.54772	ดีมาก
Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อน	5	2.00	4.00	3.2000	.83666	ดี
รวมมาตรฐานความน่าเชื่อถือของนักประเมิน	5	2.50	3.75	3.2500	.50000	ดี
Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า	5	2.00	3.00	2.4000	.54772	พอใช้
Ce5.2 การจัดสรรงบประมาณสำหรับการประเมิน	5	1.00	3.00	1.6000	.89443	ปรับปรุง
รวมมาตรฐานประสิทธิผลของการประเมิน	5	1.50	3.00	2.0000	.61237	พอใช้

สรุปผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ส่วนใหญ่อยู่ในระดับดี คือ มาตรฐานความน่าเชื่อถือของนักประเมิน (Credibility) จริยธรรมในการประเมิน (Ethicality) อรรถประโยชน์ (Utility) และความตรงของการประเมิน (Validity) ยกเว้น มาตรฐานประสิทธิผลต้นทุน (Cost-effectiveness) ที่อยู่ในระดับพอใช้ เมื่อพิจารณารายมาตรฐาน พบว่า มาตรฐานอรรถประโยชน์ มหาวิทยาลัยเทคโนโลยีราชมงคล และมหาวิทยาลัยของรัฐ อยู่ในระดับดีมาก รองลงมา คือ มหาวิทยาลัยราชภัฏและมหาวิทยาลัยเอกชน อยู่ในระดับดี มาตรฐานความน่าเชื่อถือของนักประเมิน มหาวิทยาลัยของรัฐอยู่ในระดับดีมาก รองลงมา คือ มหาวิทยาลัยเทคโนโลยีราชมงคล มหาวิทยาลัยเอกชน และมหาวิทยาลัยราชภัฏ อยู่ในระดับดีมาตรฐานจริยธรรมในการประเมิน ส่วนใหญ่อยู่ในระดับดี ได้แก่ มหาวิทยาลัยราชภัฏ มหาวิทยาลัยเอกชน และมหาวิทยาลัยของรัฐ ตามลำดับ ยกเว้น มหาวิทยาลัยเทคโนโลยีราชมงคล อยู่ในระดับดีมาก

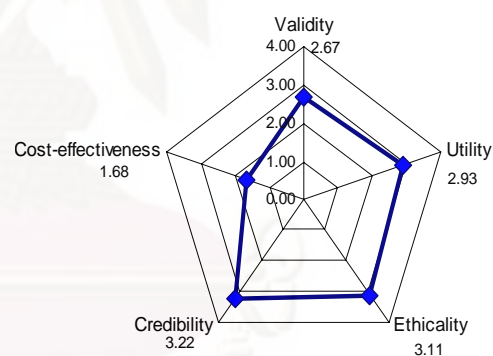
มาตรฐานความตรงของการประเมิน อยู่ในระดับดีทุกกลุ่มสถาบัน ส่วนมาตรฐานประสิทธิผลต้นทุน อยู่ในระดับพอใช้ คือ มหาวิทยาลัยเทคโนโลยีราชมงคล มหาวิทยาลัยราชภัฏ และ ตามลำดับ ยกเว้น มหาวิทยาลัยของรัฐ และมหาวิทยาลัยเอกชน อยู่ในระดับควรปรับปรุง ดังภาพที่ 4.9-4.13



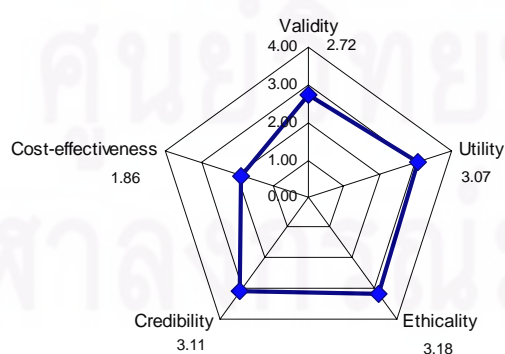
ภาพที่ 4.9 ผลการประเมินอภิมานของสถาบันอุดมศึกษา ในภาพรวม



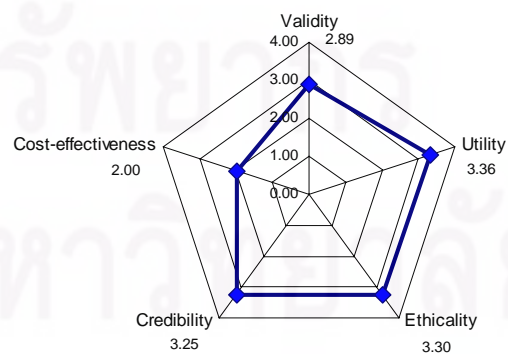
ภาพที่ 4.10 ผลการประเมินอภิมานมหาวิทยาลัยของรัฐ



ภาพที่ 4.11 ผลการประเมินอภิมานมหาวิทยาลัยเอกชน



ภาพที่ 4.12 ผลการประเมินอภิมานมหาวิทยาลัยราชภัฏ



ภาพที่ 4.13 ผลการประเมินอภิมานมหาวิทยาลัยเทคโนโลยีราชมงคล

3. ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิमान

ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันรวมทั้งหมด 50 แห่ง ในภาพรวม พบว่า ส่วนใหญ่อยู่ในระดับดี จำนวน 33 แห่ง (ร้อยละ 66.0) จำแนกเป็น มหาวิทยาลัยราชภัฏ 14 แห่ง มหาวิทยาลัยเอกชน 11 แห่ง มหาวิทยาลัยของรัฐ 5 แห่ง และมหาวิทยาลัยเทคโนโลยีราชมงคล 3 แห่ง ผลการประเมินอยู่ในระดับพอใช้ จำนวน 12 แห่ง (ร้อยละ 24.0) จำแนกเป็นมหาวิทยาลัยเอกชน 5 แห่ง มหาวิทยาลัยราชภัฏ 4 แห่ง มหาวิทยาลัยของรัฐ 2 แห่ง และมหาวิทยาลัยเทคโนโลยีราชมงคล 1 แห่ง และผลการประเมินอยู่ในระดับดีมาก จำนวน 5 แห่ง (ร้อยละ 10.0) จำแนกเป็นมหาวิทยาลัยเอกชน 3 แห่ง มหาวิทยาลัยของรัฐและมหาวิทยาลัยเทคโนโลยีราชมงคลกลุ่มละ 1 แห่ง เมื่อพิจารณารายกลุ่มสถาบัน พบว่า มหาวิทยาลัยของรัฐ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 5 แห่ง (ร้อยละ 62.5) รองลงมา อยู่ในระดับพอใช้และดีมาก จำนวน 2 แห่ง และ 1 แห่ง (ร้อยละ 25.0, 12.5) ตามลำดับ มหาวิทยาลัยเอกชน ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 11 แห่ง (ร้อยละ 57.9) รองลงมา อยู่ในระดับพอใช้และดีมาก จำนวน 5 แห่ง และ 3 แห่ง (ร้อยละ 26.3, 15.8) ตามลำดับ มหาวิทยาลัยราชภัฏ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 14 แห่ง (ร้อยละ 77.8) รองลงมา อยู่ในระดับพอใช้ จำนวน 4 แห่ง (ร้อยละ 22.2) สำหรับมหาวิทยาลัยเทคโนโลยีราชมงคล ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 3 แห่ง (ร้อยละ 60.0) รองลงมา อยู่ในระดับพอใช้และดีมาก จำนวน 1 แห่ง เท่ากัน (ร้อยละ 20.0)

ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันในมาตรฐานความตรงของการประเมิน พบว่า ส่วนใหญ่อยู่ในระดับดี จำนวน 22 แห่ง (ร้อยละ 44.0) ผลการประเมินอยู่ในระดับพอใช้ จำนวน 16 แห่ง (ร้อยละ 32.0) ผลการประเมินอยู่ในระดับดีมาก จำนวน 9 แห่ง (ร้อยละ 18.0) และผลการประเมินในระดับควรปรับปรุง จำนวน 3 แห่ง (ร้อยละ 6.0) เมื่อพิจารณารายกลุ่มสถาบัน พบว่า มหาวิทยาลัยของรัฐ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 5 แห่ง (ร้อยละ 62.5) รองลงมา อยู่ในระดับพอใช้ จำนวน 2 แห่ง (ร้อยละ 25.0) และระดับดีมาก จำนวน 1 แห่ง (ร้อยละ 12.5) มหาวิทยาลัยเอกชน ส่วนใหญ่มีผลการประเมินอยู่ในระดับพอใช้ จำนวน 8 แห่ง (ร้อยละ 42.1) รองลงมา อยู่ในระดับดี จำนวน 7 แห่ง (ร้อยละ 36.8) ระดับดีมาก 3 แห่ง และระดับควรปรับปรุง 1 แห่ง (ร้อยละ 15.8, 5.3) ตามลำดับ มหาวิทยาลัยราชภัฏ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 9 แห่ง (ร้อยละ 50.0) รองลงมา อยู่ในระดับพอใช้ จำนวน 4 แห่ง (ร้อยละ 22.2) ระดับดีมาก 3 แห่ง และระดับควรปรับปรุง 2 แห่ง (ร้อยละ 16.7, 11.1) ตามลำดับ สำหรับมหาวิทยาลัย

เทคโนโลยีราชมงคล ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก และพอใช้ จำนวน 2 แห่งเท่ากัน (ร้อยละ 40.0) รองลงมา อยู่ในระดับดี จำนวน 1 แห่ง (ร้อยละ 20.0)

ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิธานรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันในมาตรฐานอรรถประโยชน์ พบว่า ส่วนใหญ่อยู่ในระดับดี จำนวน 24 แห่ง (ร้อยละ 48.0) รองลงมาคือ ผลการประเมินอยู่ในระดับดีมาก จำนวน 19 แห่ง (ร้อยละ 38.0) ผลการประเมินอยู่ในระดับพอใช้ จำนวน 7 แห่ง (ร้อยละ 14.0) เมื่อพิจารณารายกลุ่มสถาบัน พบว่า มหาวิทยาลัยของรัฐ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 5 แห่ง (ร้อยละ 62.5) รองลงมา อยู่ในระดับดี จำนวน 3 แห่ง (ร้อยละ 37.5) มหาวิทยาลัยเอกชน ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 10 แห่ง (ร้อยละ 52.6) รองลงมา อยู่ในระดับพอใช้ จำนวน 5 แห่ง (ร้อยละ 26.3) ระดับดีมาก 4 แห่ง (ร้อยละ 21.1) มหาวิทยาลัยราชภัฏ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 9 แห่ง (ร้อยละ 50.0) รองลงมา อยู่ในระดับดีมาก จำนวน 7 แห่ง (ร้อยละ 38.9) ระดับพอใช้ 2 แห่ง (ร้อยละ 11.1) สำหรับมหาวิทยาลัยเทคโนโลยีราชมงคล ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 3 แห่ง (ร้อยละ 60.0) รองลงมา อยู่ในระดับดี จำนวน 2 แห่ง (ร้อยละ 40.0)

ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิธานรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันในมาตรฐานจริยธรรมในการประเมิน พบว่า ส่วนใหญ่อยู่ในระดับดีมาก จำนวน 27 แห่ง (ร้อยละ 54.0) รองลงมาคือ ผลการประเมินอยู่ในระดับดี จำนวน 17 แห่ง (ร้อยละ 34.0) ผลการประเมินอยู่ในระดับพอใช้และควรปรับปรุง จำนวน 3 แห่งเท่ากัน (ร้อยละ 6.0) เมื่อพิจารณารายกลุ่มสถาบัน พบว่า มหาวิทยาลัยของรัฐ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 4 แห่ง (ร้อยละ 50.0) รองลงมา อยู่ในระดับดีและควรปรับปรุง จำนวน 2 แห่ง เท่ากัน (ร้อยละ 25.0) มหาวิทยาลัยเอกชน ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 9 แห่ง (ร้อยละ 47.4) รองลงมา อยู่ในระดับดี จำนวน 8 แห่ง (ร้อยละ 42.1) ระดับพอใช้และควรปรับปรุง 1 แห่งเท่ากัน (ร้อยละ 5.3) มหาวิทยาลัยราชภัฏ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 10 แห่ง (ร้อยละ 55.6) รองลงมา อยู่ในระดับดี จำนวน 6 แห่ง (ร้อยละ 33.3) ระดับพอใช้ 2 แห่ง (ร้อยละ 11.1) สำหรับมหาวิทยาลัยเทคโนโลยีราชมงคล ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 4 แห่ง (ร้อยละ 80.0) รองลงมา อยู่ในระดับดี จำนวน 1 แห่ง (ร้อยละ 20.0)

ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันในมาตรฐานความน่าเชื่อถือของนักประเมิน พบว่า ส่วนใหญ่อยู่ในระดับดีมากและดี จำนวน 21 แห่ง เท่ากัน (ร้อยละ 42.0) รองลงมาคือ ผลการประเมินอยู่ในระดับพอใช้ จำนวน 8 แห่ง (ร้อยละ 16.0) เมื่อพิจารณารายกลุ่มสถาบัน พบว่า มหาวิทยาลัยของรัฐ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 4 แห่ง (ร้อยละ 50.0) รองลงมา อยู่ในระดับดี จำนวน 3 แห่ง (ร้อยละ 37.5) และระดับพอใช้ จำนวน 1 แห่ง (ร้อยละ 12.5) มหาวิทยาลัยเอกชน ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 9 แห่ง (ร้อยละ 47.4) รองลงมา อยู่ในระดับดี จำนวน 7 แห่ง (ร้อยละ 36.8) ระดับพอใช้ 3 แห่ง (ร้อยละ 15.8) มหาวิทยาลัยราชภัฏ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 10 แห่ง (ร้อยละ 55.6) รองลงมา อยู่ในระดับดีมาก จำนวน 5 แห่ง (ร้อยละ 27.8) ระดับพอใช้ 3 แห่ง (ร้อยละ 16.7) สำหรับมหาวิทยาลัยเทคโนโลยีราชมงคล ส่วนใหญ่มีผลการประเมินอยู่ในระดับดีมาก จำนวน 3 แห่ง (ร้อยละ 66.0) รองลงมา อยู่ในระดับดี และพอใช้ จำนวน 1 แห่ง เท่ากัน (ร้อยละ 20.0)

ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันในมาตรฐานประสิทธิภาพในการประเมิน พบว่า ส่วนใหญ่อยู่ในระดับควรปรับปรุง จำนวน 26 แห่ง (ร้อยละ 52.0) รองลงมาคือ ผลการประเมินอยู่ในระดับพอใช้ จำนวน 20 แห่ง (ร้อยละ 40.0) และระดับดี จำนวน 4 แห่ง (ร้อยละ 8.0) เมื่อพิจารณารายกลุ่มสถาบัน พบว่า มหาวิทยาลัยของรัฐ ส่วนใหญ่มีผลการประเมินอยู่ในระดับควรปรับปรุง จำนวน 4 แห่ง (ร้อยละ 50.0) รองลงมา อยู่ในระดับพอใช้ จำนวน 2 แห่ง (ร้อยละ 25.0) และระดับดี จำนวน 1 แห่ง (ร้อยละ 12.5) มหาวิทยาลัยเอกชน ส่วนใหญ่มีผลการประเมินอยู่ในระดับควรปรับปรุง จำนวน 11 แห่ง (ร้อยละ 57.9) รองลงมา อยู่ในระดับพอใช้ จำนวน 7 แห่ง (ร้อยละ 36.8) ระดับดี 1 แห่ง (ร้อยละ 5.3) มหาวิทยาลัยราชภัฏ ส่วนใหญ่มีผลการประเมินอยู่ในระดับพอใช้ จำนวน 9 แห่ง (ร้อยละ 50.0) รองลงมา อยู่ในระดับปรับปรุง จำนวน 8 แห่ง (ร้อยละ 44.4) ระดับดี 1 แห่ง (ร้อยละ 5.6) สำหรับมหาวิทยาลัยเทคโนโลยีราชมงคล ส่วนใหญ่มีผลการประเมินอยู่ในระดับพอใช้และปรับปรุง จำนวน 2 แห่ง เท่ากัน (ร้อยละ 40.0) รองลงมา อยู่ในระดับดี จำนวน 1 แห่ง (ร้อยละ 20.0) ดังตารางที่ 4.31

จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.31 ความถี่ ร้อยละของระดับผลการประเมินอภิमान จำแนกตามมาตรฐานและกลุ่มสถาบัน

มาตรฐาน	validity				utility				ethicity			
	ปรับปรุง (%)	พอใช้ (%)	ดี (%)	ดีมาก (%)	ปรับปรุง (%)	พอใช้ (%)	ดี (%)	ดีมาก (%)	ปรับปรุง (%)	พอใช้ (%)	ดี (%)	ดีมาก (%)
มหาวิทยาลัย												
ของรัฐ	0 (0.0)	2 (25.0)	5 (62.5)	1 (12.5)	0 (0.0)	0 (0.0)	3 (37.5)	5 (62.5)	2 (25.0)	0 (0.0)	2 (25.0)	4 (50.0)
เอกชน	1 (5.3)	8 (42.1)	7 (36.8)	3 (15.8)	0 (0.0)	5 (26.3)	10 (52.6)	4 (21.1)	1 (5.3)	1 (5.3)	8 (42.1)	9 (47.4)
ราชภัฏ	2 (11.1)	4 (22.2)	9 (50.0)	3 (16.7)	0 (0.0)	2 (11.1)	9 (50.0)	7 (38.9)	0 (0.0)	2 (11.1)	6 (33.3)	10 (55.6)
ราชมณฑล	0 (0.0)	2 (40.0)	1 (20.0)	2 (40.0)	0 (0.0)	0 (0.0)	2 (40.0)	3 (60.0)	0 (0.0)	0 (0.0)	1 (20.0)	4 (80.0)
รวม	3 (6.0)	16 (32.0)	22 (44.0)	9 (18.0)	0 (0.0)	7 (14.0)	24 (48.0)	19 (38.0)	3 (6.0)	3 (6.0)	17 (34.0)	27 (54.0)
มาตรฐาน	credibility				Cost-effectiveness				total			
	ปรับปรุง (%)	พอใช้ (%)	ดี (%)	ดีมาก (%)	ปรับปรุง (%)	พอใช้ (%)	ดี (%)	ดีมาก (%)	ปรับปรุง (%)	พอใช้ (%)	ดี (%)	ดีมาก (%)
มหาวิทยาลัย												
ของรัฐ	0 (0.0)	1 (12.5)	3 (37.5)	4 (50.0)	5 (62.5)	2 (25.0)	1 (12.5)	0 (0.0)	0 (0.0)	2 (25.0)	5 (62.5)	1 (12.5)
เอกชน	0 (0.0)	3 (15.8)	7 (36.8)	9 (47.4)	11 (57.9)	7 (36.8)	1 (5.3)	0 (0.0)	0 (0.0)	5 (26.3)	11 (57.9)	3 (15.8)
ราชภัฏ	0 (0.0)	3 (16.7)	10 (55.6)	5 (27.8)	8 (44.4)	9 (50.0)	1 (5.6)	0 (0.0)	0 (0.0)	4 (22.2)	14 (77.8)	0 (0.0)
ราชมณฑล	0 (0.0)	1 (20.0)	1 (20.0)	3 (60.0)	2 (40.0)	2 (40.0)	1 (20.0)	0 (0.0)	0 (0.0)	1 (20.0)	3 (60.0)	1 (20.0)
รวม	0 (0.0)	8 (16.0)	21 (42.0)	21 (42.0)	26 (52.0)	20 (40.0)	4 (8.0)	0 (0.0)	0 (0.0)	12 (24.0)	33 (66.0)	5 (10.0)

ส่วนที่ 2 ผลตรวจสอบและการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา

ผลการตรวจสอบและสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา แบ่งออกเป็น 2 ประเด็น คือ 1. ผลการตรวจสอบรายงานผลการประเมินคุณภาพภายใน และ 2. ผลการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ดังมีรายละเอียดดังต่อไปนี้

1) ผลการตรวจสอบรายงานผลการประเมินคุณภาพภายใน

สรุปการตรวจสอบรายงานผลการประเมินคุณภาพภายใน ตามที่สำนักงานคณะกรรมการการอุดมศึกษา (2552) กำหนด พบว่า รายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา มีส่วนประกอบในส่วนของจุดแข็ง ครบทุกแห่ง (100%) รองลงมาคือ ข้อเสนอแนะ (98%) และแนวปฏิบัติที่ดีหรือแนวทางเสริม (96%) ตามลำดับ ส่วนการตรวจสอบความถูกต้อง นำเชื่อถือของข้อมูลตรวจสอบพบในรายงานน้อยที่สุด (46%) เป็นที่น่าสังเกตว่าส่วนสำคัญที่ควรจะมีในรายงานผลการประเมินคุณภาพภายใน ได้แก่ บทนำ และข้อมูลพื้นฐานของสถาบัน ยังพบไม่ครบทุกแห่ง ซึ่งอาจเป็นไปได้ว่า รายงานผลการประเมินคุณภาพภายในของบางแห่งเขียนเป็นบทสรุปสั้นๆ หรือแทรกอยู่ในรายงานการประเมินตนเอง ทำให้ไม่ใส่รายละเอียดในส่วนนี้ ดังตารางที่ 4.32

ตารางที่ 4.32 รายการตรวจสอบส่วนประกอบของรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ตามที่สำนักงานคณะกรรมการการอุดมศึกษากำหนด

ที่	รายการตรวจสอบ	ม.ของรัฐ	ม.เอกชน	ม.ราชภัฏ	ม.ราช	รวม
		(n=8)	(n=19)	(n=18)	มงคล (n=5)	
1	รายนามคณะผู้ประเมิน	5 (62.5)	19 (100.0)	18 (100.0)	5 (100.0)	47 (94.0)
2	บทนำ (ประวัติ)	8 (100.0)	17 (89.5)	15 (83.3)	5 (100.0)	45 (90.0)
3	ข้อมูลพื้นฐาน	8 (100.0)	16 (84.2)	14 (77.8)	5 (100.0)	43 (86.0)
4	การวางแผนการประเมิน	4 (50.0)	18 (94.7)	15 (83.3)	4 (80.0)	41 (82.0)
5	การตรวจสอบความถูกต้องนำเชื่อถือของข้อมูล	3 (37.5)	10 (55.5)	8 (44.4)	2 (40.0)	23 (46.0)
6	ผลการประเมินรายตัวบ่งชี้ (ตาราง ป.1)	6 (75.0)	17 (89.5)	17 (94.4)	4 (80.0)	44 (80.0)
7	สรุปผล IPO (ตาราง ป. 2)	7 (87.5)	17 (89.5)	17 (94.4)	5 (100.0)	46 (92.0)
8	สรุปผลตามมาตรฐานอุดม (ตาราง ป. 3)	7 (87.5)	17 (89.5)	17 (94.4)	4 (80.0)	45 (90.0)
9	สรุปผล BSC (ตาราง ป. 4)	6 (75.0)	17 (89.5)	16 (88.9)	4 (80.0)	43 (86.0)
10	จุดแข็ง	8 (100.0)	19 (100.0)	18 (100.0)	5 (100.0)	50 (100.0)
11	แนวปฏิบัติที่ดี /แนวทางเสริม	8 (100.0)	17 (89.5)	18 (100.0)	5 (100.0)	48 (96.0)

ตารางที่ 4.32 (ต่อ)

ที่	รายการตรวจสอบ	ม.ของรัฐ	ม.เอกชน	ม.ราชภัฏ	ม.ราช มงคล	รวม
		(n=8)	(n=19)	(n=18)	(n=5)	(n=50)
12.	จุดอ่อน /จุดที่ควรพัฒนา	6 (75.0)	19 (100.0)	14 (77.8)	4 (80.0)	43 (86.0)
13.	สาเหตุของปัญหา	6 (75.0)	19 (100.0)	14 (77.8)	4 (80.0)	43 (86.0)
14.	ข้อเสนอแนะ	8 (100.0)	18 (94.7)	18 (100.0)	5 (100.0)	49 (98.0)
15.	ตารางการตรวจเยี่ยม	4 (87.5)	12 (63.1)	17 (94.4)	5 (100.0)	38 (76.0)
16.	บันทึกข้อมูลการสัมภาษณ์ หรือตรวจเยี่ยม	4 (50.0)	9 (43.4)	10 (55.5)	5 (100.0)	28 (56.0)

2) ผลการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน

ผลการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ซึ่งทำการสังเคราะห์จากรายงานผลการประเมินคุณภาพภายใน และรายงานการประเมินอภิมานของนักประเมินอภิมาน จำนวน 50 ฉบับ ซึ่งสามารถจำแนกออกเป็น 3 ประเด็น คือ จุดแข็ง จุดบกพร่อง และข้อเสนอแนะ ดังมีรายละเอียดต่อไปนี้

จุดแข็ง เนื้อหาของรายงานผลการประเมินคุณภาพภายใน มีจุดแข็งดังนี้

รายงานผลการประเมินคุณภาพภายใน มีการนำเสนอผลการประเมินได้อย่างถูกต้อง เป็นที่ยอมรับ (70.0%) โดยมีการระบุจุดเด่น จุดที่ควรพัฒนาและให้ข้อเสนอแนะที่เป็นประโยชน์ เพื่อให้เห็นทิศทางการวางแผนพัฒนาได้อย่างเหมาะสม(56.0%) เนื่องมาจากการวิเคราะห์และประมวลผลได้อย่างถูกต้องทุกตัวบ่งชี้ (54.0%) มีการกำหนดวิธีการประเมินที่สามารถวางแผนและดำเนินการประเมินได้เหมาะสมกับสภาพเป็นจริงและกำหนดช่วงเวลาในการประเมินและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม (52.0%) นอกจากนี้ รายงานการประเมินเขียนด้วยภาษาที่ชัดเจน และมีการนำเสนอด้วยสื่อที่เข้าใจง่าย ซึ่งผู้ประเมินมีความสามารถในการวางแผนและจัดการประเมินอย่างเหมาะสม โดยมีการตรวจสอบความถูกต้องของผลการประเมินจากผู้เกี่ยวข้อง และมีการระบุผู้มีส่วนได้ส่วนเสียในการประเมิน การจัดทำรายงานผลการประเมินได้เสร็จทันเวลา ซึ่งผู้ประเมินสามารถเข้าใจบริบทและให้ข้อเสนอแนะที่นำไปสู่เป้าหมายที่ต้องการได้และมีจริยธรรม ทำการประเมินรายงานได้นำเชื่อถือได้รับการยอมรับจากผู้ที่เกี่ยวข้องในการประเมินได้ อีกทั้งมีการตรวจสอบความน่าเชื่อถือของแหล่งข้อมูล วิธีการเก็บข้อมูล และผู้เก็บรวบรวมข้อมูล รวมถึงมีการเปิดเผยข้อจำกัดที่มีผลกระทบต่อผลการประเมิน

(26.0-48.0%) แต่รายงานผลการประเมินส่วนน้อย ยังพบว่า การเขียนรายงานเป็นไปตามรูปแบบ สกอ. เนื่องจาก สกอ.มีเกณฑ์และมาตรฐานในการประเมินที่มีการปรับปรุงพัฒนาอย่างต่อเนื่อง ทั้งนี้รายงานมีการระบุวัตถุประสงค์และวัตถุประสงค์หรือเป้าหมายของการประเมิน และมีการ นำเสนอข้อเสนอแนะเพิ่มเติม และข้อเสนอแนะในการเขียนรายงานอีกด้วย (2.0-24.0%)

จุดบกพร่อง เนื้อหาของรายงานผลการประเมินคุณภาพภายใน มีจุดบกพร่องดังนี้ รายงานผลการประเมินคุณภาพภายในส่วนใหญ่ พบว่า ไม่ได้ระบุหลักฐานที่ แสดงการวางแผนเกี่ยวกับงบประมาณ (76.0%) ไม่มีรายละเอียดการวางแผนจัดการด้าน ทรัพยากรในการประเมิน (54.0%) ขาดการวิเคราะห์ผลการประกันคุณภาพในอดีตซึ่งจะทำให้เกิด การปรับปรุงอย่างต่อเนื่อง (36.0%) ขาดการนำเสนอวัตถุประสงค์ของการประเมิน (34.0%) และ ขาดเอกสารหลักฐานที่ตกลงทำการประเมินระหว่างสถาบัน (32.0%) นอกจากนี้ยังพบว่ารายงาน ผลการประเมินบางส่วน รูปแบบของรายงานการประเมินคุณภาพภายในของคณะกรรมการ ฯ ไม่มี การแสดงถึงรายละเอียดชัดเจน ไม่ครบถ้วน ไม่มีการนำเสนอร่องรอย /หลักฐาน/ รายละเอียดการ สัมภาษณ์ และการบันทึกข้อมูลภาคสนามที่พบและนำมาใช้ประกอบการพิจารณาตัดสินผลการ ประเมินแต่ละตัวบ่งชี้ ไม่ปรากฏหลักฐานที่เกี่ยวข้องกับการวางแผนการประเมินและผู้มีส่วนได้ ส่วนเสียไว้ และยังขาดการตรวจสอบความถูกต้องของข้อมูล ขาดข้อมูลพื้นฐาน ภูมิหลังของ สถาบัน รวมถึงไม่มีการระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม (22.0- 30.0%) แต่ทั้งนี้รายงานผลการประเมินส่วนน้อย พบว่า ยังมีข้อผิดพลาดในการวิเคราะห์ข้อมูลใน บางตัวบ่งชี้ ข้อมูลในรายงานไม่ตรงกับข้อมูลพื้นฐาน (common data set) ระบุจุดแข็งจุดอ่อนไว้ ในบางองค์ประกอบไม่ครบถ้วน รายงานยังขาดในส่วนของบทสรุปผู้บริหาร และการสรุปผล ประเมินในแต่มาตรฐานในเชิงการบรรยาย มิได้ให้ข้อเสนอแนะเกี่ยวกับจุดที่ควรพัฒนาบาง องค์ประกอบซึ่งจะเป็นประโยชน์ต่อสถาบัน และยังขาดการระบุวัตถุประสงค์ของสถาบัน ไม่ระบุ รายละเอียดการรับฟังข้อมูลย้อนกลับจากการใช้ผลการประเมิน นอกจากนี้ยังขาดประเด็นการ เสนอแนะการเขียนรายงานการประเมินตนเองของสถาบันให้อยู่ในรูปแบบที่สอดคล้องตามเกณฑ์ ของสกอ. (2.0-12.0%)

ข้อเสนอแนะ

ข้อเสนอแนะจากการประเมินคุณภาพภายใน ระดับอุดมศึกษา มีรายละเอียด ดังต่อไปนี้

การประเมินคุณภาพภายใน และการรายงานผลการประเมินคุณภาพภายใน ควร ปรับปรุงพัฒนารูปแบบรายงานการประเมินที่ครอบคลุม ครบถ้วน มีสารสนเทศที่เพียงพอต่อการ

นำเสนอให้สถาบันอุดมศึกษาและผู้เกี่ยวข้องตลอดจนสาธารณชนได้เข้าใจมากขึ้น (42.0%) ข้อเสนอหรือจุดที่ควรพัฒนาควรระบุสาเหตุของปัญหาและข้อเสนอแนะอย่างชัดเจน (24.0%) ควรปรับปรุงคุณภาพของผู้ประเมิน ให้มีความรู้ ความสามารถ และทักษะในการประเมินที่ถูกต้องและได้มาตรฐานที่ใกล้เคียง เพื่อประโยชน์ของคุณภาพการศึกษาต่อไป (20.0%) การประเมินในครั้งนี้อาจไม่สามารถประเมินได้โดยตรงจากผู้ประเมิน แต่ดูจากเอกสารรายงานอย่างเดียว ทำให้อาจประเมินได้ไม่ชัดเจนตามความเป็นจริง และควรมีการเปิดเผยรายงานผลการประเมินให้ครบถ้วน และเผยแพร่เอกสารการประเมินให้สาธารณชนรับทราบ (18.0%) ทั้งนี้ควรจัดหาผู้ประเมินที่ไม่มีอิทธิพลเกี่ยวกับผลประโยชน์จากกลุ่มใดที่ส่งผลต่อการประเมิน (16.0%) นอกจากนี้ การประเมินในครั้งนี้นี้ มีบุคคลภายในเป็นผู้ประเมินด้วย ดังนั้นควรทำการตรวจสอบคุณภาพของงานประเมินด้วย เพื่อขจัดความลำเอียงในการประเมิน โดยควรทำการประเมินอภิมานเนื่องจากมีประโยชน์โดยตรงต่อผู้ประเมินในการพัฒนาปรับปรุงทักษะต่าง ๆ ในการประเมินให้มีความสมบูรณ์ยิ่งขึ้นต่อไป อีกทั้งควรระบุจุดเน้นของสถาบันไว้ในรายงาน และข้อเสนอแนะ แนวทางพัฒนา โดยระบุเป็นแนวทางในระยะสั้น (ดำเนินการเร่งด่วน) ระยะกลาง (1 – 3 ปี) และระยะยาว (5 ปี) (6.0-10.0%) ดังมีรายละเอียดตารางที่ 4.33

ตารางที่ 4.33 การสังเคราะห์เนื้อหาของรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา

ที่	รายละเอียดเนื้อหา	มหาวิทยาลัย				รวม (%)
		ของรัฐ	เอกชน	ราชภัฏ	ราชมงคล	
	จุดแข็ง					
1	ผู้ประเมินมีความสามารถในการวางแผนและจัดการประเมินอย่างเหมาะสม	5	7	9	2	23 (46.0)
2	ผู้ประเมินเข้าใจบริบทและสามารถให้ข้อเสนอแนะที่นำไปสู่เป้าหมายที่ต้องการได้	3	2	9		14 (28.0)
3	ผู้ประเมินมีจริยธรรม ทำการประเมินรายงานได้นำเชื่อถือ		4	6	3	13 (26.0)
4	ผู้ประเมินมีทักษะในการสื่อสารรายงานผลการประเมินได้อย่างละเอียด เข้าใจง่าย และ	4	8	13	2	27 (54.0)
5	ผู้ประเมินมีความยุติธรรมและมีทักษะในการให้ข้อเสนอแนะที่เกิดประโยชน์และได้รับการยอมรับจากผู้ที่เกี่ยวข้องในการประเมิน	2	10	13	3	28 (56.0)
6	ผู้ประเมินได้รับการยอมรับจากผู้ที่เกี่ยวข้องในการประเมินได้	2	4	10	1	17 (34.0)
7	การกำหนดวิธีการประเมินที่สามารถวางแผนและดำเนินการประเมินได้เหมาะสมกับสภาพเป็นจริง	3	9	11	3	26 (52.0)
8	มีการระบุข้อดีของสถาบัน	3	3	1	2	9 (18.0)

ตารางที่ 4.33 (ต่อ)

ที่	รายละเอียดเนื้อหา	มหาวิทยาลัย				รวม (%)
		ของรัฐ	เอกชน	ราชภัฏ	ราชมงคล	
9	มีระบุจุดเด่น จุดที่ควรพัฒนาและให้ข้อเสนอแนะที่เป็นประโยชน์ทำให้เห็นทิศทางการวางแผนพัฒนาได้อย่างเหมาะสม	7	9	11	1	28 (56.0)
10	นำเสนอผลการประเมินอย่างถูกต้อง เป็นที่ยอมรับ	6	10	15	4	35 (70.0)
11	สกอ.มีเกณฑ์และมาตรฐานในการประเมินที่มีการปรับปรุงพัฒนาอย่างต่อเนื่อง	1	3	1	2	7 (14.0)
12	วิเคราะห์และประมวลผลได้อย่างถูกต้อง	5	9	10	3	27 (54.0)
13	มีการตรวจสอบความน่าเชื่อถือของแหล่งข้อมูล วิธีการเก็บข้อมูล และผู้เก็บรวบรวมข้อมูล	3	5	8	1	17 (34.0)
14	มีการตรวจสอบความถูกต้องของผลการประเมินจากผู้เกี่ยวข้อง	4	6	10	2	22 (44.0)
15	มีการระบุผู้มีส่วนได้ส่วนเสียในการประเมิน	4	6	10	1	21 (42.0)
16	ระบุช่วงเวลาในการประเมินและระยะเวลาในการเก็บรวบรวมข้อมูล	5	8	10	3	26 (52.0)
17	รายงานการประเมินเขียนด้วยภาษาที่ชัดเจน และมีการนำเสนอด้วยสื่อที่เข้าใจง่าย	5	6	10	3	24 (48.0)
18	มีการเปิดเผยข้อจำกัดที่มีผลกระทบต่อผลการประเมิน	2	2	9	1	14 (28.0)
19	การจัดทำรายงานผลการประเมินเสร็จทันเวลา	3	4	11	1	19 (38.0)
20	การเขียนรายงานเป็นไปตามรูปแบบ สกอ.	2		8	2	12 (24.0)
21	มีการระบุถึงวัตถุประสงค์หรือเป้าหมายของการประเมิน	3		5		8 (16.0)
22	มีการนำเสนอข้อเสนอแนะเพิ่มเติม และข้อเสนอแนะในการเขียนรายงาน	1				1 (2.0)
จุดบกพร่อง						
1	รูปแบบและรายละเอียดของรายงานผลการประเมินคุณภาพภายในของคณะกรรมากร ฯ ไม่ครบถ้วนชัดเจน	4	6	3	2	15 (30.0)
2	ขาดประเด็นการเสนอแนะการเขียนรายงานการประเมินตนเองของสถาบันให้อยู่ในรูปแบบที่สอดคล้องตามเกณฑ์ของสกอ.	1		2		3 (6.0)
3	ขาดการนำเสนอวัตถุประสงค์ของการประเมิน	2	6	8	1	17 (34.0)
4	ขาดข้อมูลพื้นฐาน ภูมิหลังของสถาบัน	2	3	6		11 (22.0)

ตารางที่ 4.33 (ต่อ)

ที่	รายละเอียดเนื้อหา	มหาวิทยาลัย				รวม (%)
		ของรัฐ	เอกชน	ราชภัฏ	ราชมงคล	
5	ขาดการตรวจสอบความถูกต้องของข้อมูล	2	5	6		13 (26.0)
6	ไม่มีการระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม	1	7	3		11 (22.0)
7	ไม่ปรากฏหลักฐานที่เกี่ยวข้องกับการวางแผนการประเมิน ผู้มีส่วนได้ส่วนเสีย ไว้ล่วงหน้า	4	4	4	1	13 (26.0)
8	มีข้อผิดพลาดในการวิเคราะห์ข้อมูลในบางตัวบ่งชี้	2	2	1		5 (10.0)
9	ข้อมูลในรายงานไม่ตรงกับข้อมูลพื้นฐาน (common data set)		1			1 (2.0)
10	ขาดการวิเคราะห์ผลการประกันคุณภาพในอดีตซึ่งจะทำให้เกิดการปรับปรุงอย่างต่อเนื่อง	3	6	9		18 (36.0)
11	ไม่ได้ระบุหลักฐานที่แสดงการวางแผนเกี่ยวกับงบประมาณ	6	14	13	5	38 (76.0)
12	ไม่มีรายละเอียดการวางแผนจัดการด้านทรัพยากรในการประเมิน	6	9	10	2	27 (54.0)
13	ระบุจุดแข็งจุดอ่อนไว้ในบางองค์ประกอบไม่ครบถ้วน	1	2	2		5 (10.0)
14	ไม่มีการนำเสนอร่องรอย /หลักฐาน/ รายละเอียดการสัมภาษณ์ และการบันทึกข้อมูลภาคสนามที่พบและนำมาใช้ประกอบการพิจารณาตัดสินผลการประเมินแต่ละตัวบ่งชี้	3	5	4	1	13 (26.0)
15	ขาดเอกสาร หลักฐานที่ตกลงทำการประเมินระหว่างสถาบัน	5	6	4	1	16 (32.0)
16	รูปแบบของรายงานการประเมินไม่มีรูปแบบที่แสดงถึงรายละเอียดชัดเจน	2	7	3	1	13 (26.0)
17	รายงานยังขาดในส่วนของบทสรุปผู้บริหาร และการสรุปผลประเมินในแต่มาตรฐานในเชิงการบรรยาย		3	1		4 (8.0)
18	การเขียนรายงานการประเมินนั้นมีการพิมพ์และจัดหน้าผิดพลาดเล็กน้อย	1			1	2 (4.0)
19	ผู้ประเมินมิได้ให้ข้อเสนอแนะเกี่ยวกับจุดที่ควรพัฒนาบางองค์ประกอบซึ่งจะเป็นประโยชน์ต่อสถาบัน			4	1	5 (10.0)
20	ขาดการระบุอัตลักษณ์ของสถาบัน	2		4		6 (12.0)
21	ไม่ระบุการรับฟังข้อมูลย้อนกลับจากการใช้ผลการประเมิน	1		2		3 (6.0)

ตารางที่ 4.33 (ต่อ)

ที่	รายละเอียดเนื้อหา	มหาวิทยาลัย				รวม (%)
		ของรัฐ	เอกชน	ราชภัฏ	ราชมงคล	
ข้อเสนอแนะ						
1	ควรปรับปรุงพัฒนารูปแบบรายงานการประเมินที่ครอบคลุม ครบถ้วน มีสารสนเทศที่เพียงพอต่อการนำเสนอให้สถานศึกษาและผู้เกี่ยวข้อง ตลอดจนสาธารณชนได้เข้าใจมากขึ้น	3	9	7	2	21 (42.0)
2	ควรจัดหาผู้ประเมินที่ไม่มีอิทธิพลเกี่ยวกับผลประโยชน์จากกลุ่มใดที่ส่งผลกระทบต่อประเมิน	1	4	2	1	8 (16.0)
3	ควรมีการเปิดเผยรายงานผลการประเมินให้ครบถ้วน และเผยแพร่เอกสารการประเมินให้สาธารณชนรับทราบ	1		6	2	9 (18.0)
4	การประเมินในครั้งนี้ มีบุคคลภายในเป็นผู้ประเมินด้วย ดังนั้นควรตรวจสอบคุณภาพของงานประเมินด้วย เพื่อจัดความลำเอียงในการประเมิน	1		3	1	5 (10.0)
5	ควรปรับปรุงคุณภาพของผู้ประเมิน ให้มีความรู้ความสามารถ และทักษะในการประเมินที่ถูกต้อง และได้มาตรฐานที่ใกล้เคียง เพื่อประโยชน์ของคุณภาพการศึกษาต่อไป	2	1	5	2	10 (20.0)
6	ข้อเสนอหรือจุดที่ควรพัฒนาควรระบุสาเหตุของปัญหาและข้อเสนอแนะอย่างชัดเจน		4	6	2	12 (24.0)
7	ข้อเสนอแนะและแนวทางพัฒนา ควรระบุเป็นแนวทางในระยะสั้น (ดำเนินการเร่งด่วน ระยะกลาง (1 – 3 ปี) และระยะยาว (5 ปี))		2	1		3 (6.0)
8	การประเมินในครั้งนี้ไม่สามารถประเมินได้โดยตรงจากผู้ประเมิน แต่ดูจากเอกสารรายงานอย่างเดียว ทำให้อาจประเมินได้ไม่ชัดเจนตามความเป็นจริง		5	4		9 (18.0)
9	ควรระบุจุดเน้นของสถาบันไว้ในรายงาน		3	1		4 (8.0)
10	การประเมินอภิมานประโยชน์โดยตรงต่อผู้ประเมินในการพัฒนาปรับปรุงทักษะต่าง ๆ ในการประเมิน ให้มีความสมบูรณ์ยิ่งขึ้นต่อไป		2	2		4 (8.0)

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยครั้งนี้มีวัตถุประสงค์ 1) เพื่อพัฒนามาตรฐานการประเมินอภิमानสำหรับประเมินรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษาในประเทศไทย 2) เพื่อตรวจสอบคุณภาพของมาตรฐานการประเมินอภิमान ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายในของสถานศึกษา ระดับอุดมศึกษา 3) เพื่อประยุกต์ใช้มาตรฐานการประเมินอภิमान ในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ของสถานศึกษา ระดับอุดมศึกษา

การวิจัยในครั้งนี้เป็นการวิจัยที่มุ่งประเมินอภิमान โดยมีขั้นตอนการดำเนินการวิจัยดังต่อไปนี้ คือ 1) ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องเพื่อพัฒนามาตรฐานการประเมินอภิमानจากมาตรฐานการประเมิน program evaluation standards และมาตรฐานการประเมินที่เกี่ยวข้อง 2) พิจารณา มาตรฐานการประเมินอภิमानที่พัฒนาขึ้น โดยตรวจสอบผลการสังเคราะห์ มาตรฐานการประเมินอภิमान ให้สอดคล้องกับการประเมินรายงานผลการประเมินคุณภาพภายใน และคู่มือการประเมินคุณภาพภายในสถานศึกษา ระดับอุดมศึกษา 3) ดำเนินการพัฒนาเครื่องมือที่ใช้ในการประเมินอภิमान คือ แบบประเมินอภิमान หลักสูตรฝึกอบรมนักประเมินอภิमान และคู่มือการประเมินอภิमान 4) ให้ผู้เชี่ยวชาญจำนวน 5 ท่าน พิจารณาตรวจสอบคุณภาพแบบประเมินอภิमान ด้านความตรงตามเนื้อหา (content validity) และดำเนินการปรับปรุงเครื่องมือตามข้อเสนอแนะของผู้เชี่ยวชาญให้เหมาะสมในการประเมิน 5) ดำเนินการออกแบบเพื่อพัฒนาหลักสูตรในการฝึกอบรมนักประเมินอภิमान ตามมาตรฐานการประเมินอภิमानที่พัฒนาขึ้น 6) คัดเลือกนักประเมินอภิमान จากผู้ที่มีความรู้ ประสบการณ์ในการประเมินคุณภาพภายใน ระดับอุดมศึกษา หรือผ่านการอบรมหลักสูตรผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษา จำนวน 28 คน รวม 2 รุ่น 7) นำนักประเมินอภิमानที่คัดเลือกมาฝึกอบรมในหลักสูตร “การพัฒนานักประเมินอภิमान” 8) ให้ผู้เข้ารับการอบรมฝึกประเมินอภิमान โดยผู้เข้าอบรมทั้ง 2 รุ่น ประเมินอภิमानรายงานผลการประเมินคุณภาพของสถาบันอุดมศึกษา รวม 5 ฉบับ ตามขั้นตอนการฝึกอบรมนักประเมินอภิमान 9) การตรวจสอบคุณภาพของแบบประเมินอภิमानด้านความเที่ยง (reliability) โดยให้นักประเมินอภิमानประเมินรายงานผลการประเมินคุณภาพภายในไขว้กัน 2 คนต่อ 1 ฉบับ เพื่อวิเคราะห์หาค่าความเที่ยงแบบสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) 10) นำผลการตรวจสอบคุณภาพ

มาตรฐานการประเมินอภิมานและเครื่องมือ ที่ผ่านการทดลองประเมิน มาพิจารณาปรับปรุงเครื่องมือในการประเมินอภิมานให้เหมาะสม 11) นำรายงานผลการประเมินคุณภาพภายใน จำนวน 10 ฉบับไปให้ผู้ทรงคุณวุฒิด้านการประกันคุณภาพจำนวน 10 ท่าน ประเมินคุณภาพรายงานฯ เพื่อบ่งชี้ระดับคุณภาพ เป็นการพิจารณาคุณภาพด้านความตรงเชิงโครงสร้าง (construct validity) ด้วยวิธีเปรียบเทียบกับกลุ่มรู้ชัด (known group method) เพื่อนำมาเปรียบเทียบกับผลการประเมินอภิมานของนักประเมินอภิมาน 12) นักประเมินอภิมานที่ผ่านการอบรมในหลักสูตร การฝึกอบรมพัฒนา นักประเมินอภิมาน ไปดำเนินการประเมินอภิมานรายงานผลการประเมินคุณภาพภายในระดับอุดมศึกษา จำนวน 45 แห่ง ตามขั้นตอน กระบวนการที่ได้รับการฝึกอบรม เพื่อพิจารณาหาคุณภาพของรายงานผลการประเมินคุณภาพภายใน ตามมาตรฐานการประเมินอภิมานที่พัฒนา 13) นำผลการประเมินอภิมานที่ได้จากการประเมินไปวิเคราะห์ข้อค้นพบ เชิงปริมาณ และเชิงคุณภาพ ดำเนินการสรุปผลข้อค้นพบและเขียนรายงาน

สรุปผลการวิจัย

ผู้วิจัยจึงขอสรุปผลการวิจัยเป็น 3 ตอน ดังนี้ 1) การพัฒนามาตรฐานการประเมินอภิมาน ระดับอุดมศึกษา 2) การตรวจสอบคุณภาพของมาตรฐานการประเมินอภิมาน 3) การประยุกต์ใช้มาตรฐานการประเมินอภิมานในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา

ตอนที่ 1 การพัฒนามาตรฐานการประเมินอภิมาน ระดับอุดมศึกษา

มาตรฐานการประเมินอภิมาน และตัวบ่งชี้ที่พัฒนาขึ้นได้มาจากการสังเคราะห์มาตรฐานการประเมินใน 7 แนวทางประเมิน ได้แก่ 1. Program Evaluation Standards (1994), 2. The African Evaluation guidelines (2002), 3. The DeGEval's Standards (2003), 4. SEVAL's Evaluation Standards (2004), 5. Guiding Principles for Evaluators (2003), 6. Essential Skills Series in Evaluation (1999), 7. Essential Competencies for Program Evaluators (2006) และการประเมินอภิมานของ สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) (สมศ.) พบว่า มีความสอดคล้องกับมาตรฐานการประเมินอภิมานใน 5 มาตรฐาน 38 ตัวบ่งชี้ โดยมีคำอธิบายรายละเอียดและข้อมูลดังต่อไปนี้

มาตรฐานที่ 1 ความตรงของการประเมิน (Validity) หมายถึง คุณลักษณะของงานประเมิน ที่มีการจัดการเอกสารในการประเมิน การวิเคราะห์บริบท กำหนดวัตถุประสงค์และออกแบบการประเมิน การเก็บรวบรวมสารสนเทศ วิเคราะห์ แปลผลและสรุปสารสนเทศอย่างมี

คุณภาพ และตรวจสอบการประเมินอย่างถูกต้องสมเหตุสมผล ประกอบด้วย 13 ตัวบ่งชี้ คือ Va1.1 การระบุบริบทของสถาบันที่ส่งผลกระทบต่อการศึกษาประเมิน, Va1.2 การระบุถึงอัตลักษณ์ของสถาบัน, Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมินอย่างชัดเจน, Va1.4 การออกแบบการประเมินที่สอดคล้องกับสภาพ และบริบทของสถาบัน, Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสารในการประเมิน, Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ, Va1.7 การตรวจสอบความถูกต้องของข้อมูล, Va1.8 การตรวจสอบคุณภาพของสารสนเทศ ตรงกับความต้องการของผู้ใช้ข้อมูล, Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูลอย่างถูกต้องตามวิธีการที่กำหนด, Va1.10 การแปลผลและสรุปผลการประเมินถูกต้องตามหลักฐานที่ปรากฏ, Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ, Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง, และ Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมินจากผู้ที่เกี่ยวข้อง

มาตรฐานที่ 2 อรรถประโยชน์ (Utility) หมายถึง คุณลักษณะของงานประเมินที่เป็นประโยชน์ต่อผู้มีส่วนได้ส่วนเสีย และผู้ที่เกี่ยวข้อง มีการตัดสินใจคุณค่า และรายงานผลการประเมินอย่างชัดเจน มีการเผยแพร่ผลการประเมินทันความต้องการของผู้ใช้ผลการประเมิน และเป็นประโยชน์ในการวางแผน ปรับปรุงการดำเนินงานในทิศทางที่เสนอแนะ ประกอบด้วย 10 ตัวบ่งชี้ คือ Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องในการประเมิน, Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม, Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล โดยคำนึงถึงข้อจำกัดของทรัพยากรการประเมิน, Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมินสอดคล้องกับสภาพจริง, Ut 2.5 ผลการประเมินเป็นประโยชน์ ซึ่งให้เห็นทิศทางการปรับปรุง การพัฒนาการดำเนินงานของสถาบัน, Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด, Ut 2.7 รายงานผลการประเมินครบถ้วน ครอบคลุมผลการประเมิน และข้อเสนอแนะเพื่อการพัฒนา, Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ และมีสื่อประกอบ, Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา, และ Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม

มาตรฐานที่ 3 จริยธรรมในการประเมิน (Ethicality) หมายถึง คุณลักษณะของงานประเมินที่มีการกำหนดวิธีการประเมินได้เหมาะสมกับสภาพจริง ได้รับการยอมรับ มีการคำนึงถึงผู้ที่เกี่ยวข้องหลายกลุ่ม มีการปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงสิทธิมนุษยชน และประโยชน์สาธารณะ ทำการประเมินที่มีความถูกต้องยุติธรรมต่อผู้รับผลการประเมิน รวมถึงเปิดเผยผลการประเมินแก่ผู้ที่ได้รับผลกระทบจากการประเมิน ประกอบด้วย 9 ตัวบ่งชี้ คือ Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมินตลอดช่วงเวลาประเมิน, Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง และปราศจากอคติการประเมิน, Et 3.3 การปรับปรุงคุณภาพการประเมินอย่าง

ต่อเนื่อง โดยคำนึงถึงความต้องการจำเป็นของผู้รับการประเมิน, Et 3.4 การจัดทำเอกสารข้อตกลง การประเมิน, Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัดที่มีผลกระทบต่อ การประเมิน, Et 3.6 การกำหนดวิธีการประเมินโดยคำนึงถึงความปลอดภัย รักษาความลับ และไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล, Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง โดยคำนึงถึงความแตกต่าง หลากหลาย ของผู้ที่เกี่ยวข้อง, Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย, และ Et 3.9 การประเมินที่ ครบถ้วน ชัดเจนสอดคล้องกับมาตรฐานและ ตัวบ่งชี้

มาตรฐานที่ 4 ความน่าเชื่อถือของนักประเมิน (Credibility) หมายถึงคุณลักษณะ ของงานประเมินที่ได้มาจากนักประเมินที่มีสมรรถนะในการประเมิน ได้รับการยอมรับ และไม่มี ผลประโยชน์ทับซ้อนเพื่อสร้างความน่าเชื่อถือของข้อค้นพบและสารสนเทศที่ได้รับจากการประเมิน ประกอบด้วย 4 ตัวบ่งชี้ คือ Cr4.1 ผู้ประเมินมีความสามารถ ในการประเมิน, Cr 4.2 ผู้ประเมินมี ทักษะในการสื่อสารผลการประเมิน, Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมินอย่าง เหมาะสม, Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อนของผู้ประเมิน

มาตรฐานที่ 5 ประสิทธิภาพต้นทุน (Cost-effectiveness) หมายถึงคุณลักษณะของ งานประเมิน ที่คำนึงถึงความคุ้มค่าของทรัพยากรที่จำเป็นในการประเมิน มีค่าใช้จ่ายในการ ประเมินอย่างสมเหตุสมผล มีหลักฐานชัดเจนและมีความโปร่งใสตรวจสอบได้ของต้นทุนการ ประเมิน ประกอบด้วย 2 ตัวบ่งชี้ คือ Ce5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า และ Ce5.2 การจัดสรรงบประมาณสำหรับการประเมินอย่างเหมาะสม

การนำมาตรฐานการประเมินอภิมาน ไปวิเคราะห์หาความตรงเชิงโครงสร้าง ด้วย วิธีวิเคราะห์องค์ประกอบเชิงยืนยัน จากข้อมูลเชิงประจักษ์ พบว่า มาตรฐานการประเมินอภิมาน (META-EVALUATION) มีความสอดคล้องกลมกลืนกับข้อมูลเชิงประจักษ์ พิจารณาจากค่า ไคสแควร์ที่แตกต่างจากศูนย์อย่างไม่มีนัยสำคัญทางสถิติ ($\chi^2 = 0.84, df = 3, p = 0.84067$) ค่า ดัชนีวัดระดับความกลมกลืน (GFI) มีค่าเท่ากับ 0.99 ค่าดัชนีวัดระดับความกลมกลืนที่ปรับแก้แล้ว (AGFI) มีค่าเท่ากับ 0.97 และค่าดัชนีรากของค่าเฉลี่ยกำลังสองของส่วนที่เหลือ (RMSEA) มีค่า เท่ากับ 0.000 โดยมาตรฐานที่มีน้ำหนักความสำคัญมากที่สุด คือ มาตรฐานอรรถประโยชน์ (utility) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.90 และมีการแปรผันร่วมกับมาตรฐานการประเมินอภิมาน ร้อยละ 81 รองลงมาคือ มาตรฐานความตรงของการประเมิน (validity) ความน่าเชื่อถือของ นักประเมิน (credibility) และจริยธรรมในการประเมิน (ethicality) มีค่าน้ำหนักองค์ประกอบ เท่ากับ 0.86, 0.82 และ 0.40 ตามลำดับ

ตอนที่ 2 การตรวจสอบคุณภาพของมาตรฐานการประเมินอภิमान

ผลการตรวจสอบคุณภาพของมาตรฐานประเมินอภิमान แบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 ผลการตรวจสอบคุณภาพด้านความตรงเชิงโครงสร้าง ส่วนที่ 2 ผลการตรวจสอบคุณภาพด้านความเที่ยง ดังมีรายละเอียดดังต่อไปนี้

ส่วนที่ 1 ผลการตรวจสอบคุณภาพ ด้านความตรงเชิงโครงสร้าง (construct validity) โดยใช้เทคนิคกลุ่มรู้ชัด (known group technique) โดยใช้ผู้ทรงคุณวุฒิด้านการประกันคุณภาพการศึกษา จำนวน 10 ท่าน เป็นผู้บ่งชี้คุณภาพของรายงานผลการประเมินคุณภาพภายใน โดยสามารถจำแนกรายงานผลการประเมินออกเป็น 2 กลุ่ม กลุ่มสูง 6 ฉบับ และกลุ่มต่ำ 4 ฉบับ คือ กลุ่มแรก เป็นรายงานที่ผลการประเมินคุณภาพสูงในระดับดีขึ้นไป และกลุ่มที่สอง เป็นรายงานที่มีผลการประเมินคุณภาพต่ำกว่าระดับดี เพื่อเปรียบเทียบคุณภาพกับผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ที่ประเมินโดยนักประเมินอภิमानที่ผ่านการฝึกอบรมตามหลักสูตรที่พัฒนาขึ้น พบว่า ผลการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ระหว่างนักประเมินอภิमानประเมินรายงานตามกลุ่มที่แบ่งโดยผู้ทรงคุณวุฒิบ่งชี้ กลุ่มคุณภาพทั้งสองกลุ่ม แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ($t = 2.621$, $p\text{-value} = .031$)

ส่วนที่ 2 ผลการตรวจสอบคุณภาพมาตรฐานการประเมินอภิमानด้านความเที่ยง (reliability) ระหว่างนักประเมินอภิमानต่างทีมกัน

ผลการประมาณค่าความแปรปรวนของแบบประเมินอภิमानที่มีผลต่อค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) จากการประเมินของนักประเมินอภิमान จำนวน 1, 2, และ 3 คน ที่ทำการประเมินรายงานผลการประเมินคุณภาพภายใน จำนวน 3, 5, 7, และ 9 ฉบับ โดยประเมินไขว้กัน โดยการวิธีการหาความเที่ยงโดยวิธีสัมประสิทธิ์การสรุปอ้างอิง (G-Theory) I X (R:P) Design เมื่อทำการประเมินด้วยแบบประเมินอภิमानชุดเดียวกันจำนวน 38 ตัวบ่งชี้ สัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์ (phi coefficient) ในกรณีใช้ รายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา จำนวน 3, 5, 7, และ 9 ฉบับ และจำนวนที่ให้นักประเมินอภิमानจำนวน 1, 2, 3 คน กรณีใช้รายงานผลประเมินคุณภาพภายใน 3 ฉบับ มีค่า 0.542, 0.689, และ 0.758 ตามลำดับ หากเพิ่มรายงานผลประเมินคุณภาพภายในเท่ากับ 5 ฉบับ มีค่า 0.663, 0.787, และ 0.839 ตามลำดับ กรณีเพิ่มรายงานผลประเมินคุณภาพภายในเท่ากับ 7 ฉบับ มีค่า 0.734, 0.838, และ 0.880 ตามลำดับ และกรณีเพิ่มรายงานผลประเมินคุณภาพภายในเป็น 9 ฉบับ มีค่า 0.780, 0.869, และ 0.904 ตามลำดับ แสดงให้เห็นว่ากรณี เพิ่มจำนวนรายงานผลการประเมินคุณภาพภายใน และเพิ่มจำนวนนักประเมินอภิमान ค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) มีค่ามากขึ้น จากผลการ

วิเคราะห์พบว่า ควรกำหนดนักประเมินอภิमानจำนวน 3 คนขึ้นไป กรณีประเมินรายงานผลการประเมินฯ จำนวน 3 ฉบับ และควรกำหนดนักประเมินอภิमानจำนวน 2 คนขึ้นไป สำหรับประเมินรายงานผลการประเมินฯ จำนวน 5 ฉบับ และนอกจากนี้ควรใช้นักประเมินอภิमानตั้งแต่ 1 คนขึ้นไป สำหรับประเมินรายงานฯ จำนวน 7 และ 9 ฉบับ ตามลำดับ

ตอนที่ 3 การประยุกต์ใช้มาตรฐานการประเมินอภิमानในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา

ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ผู้วิจัยขอนำเสนอแบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา และส่วนที่ 2 ผลการตรวจสอบและสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ดังมีรายละเอียดดังต่อไปนี้

ส่วนที่ 1 ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา

ผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายในของสถาบันอุดมศึกษา ผู้วิจัยขอนำเสนอผลการประเมินอภิमानใน 3 หัวข้อ คือ 1. ผลการประเมินอภิमानของสถาบันอุดมศึกษาในภาพรวม 2. ผลการประเมินอภิमानของสถาบันอุดมศึกษาจำแนกตามกลุ่มสถาบัน 4 กลุ่ม ได้แก่ มหาวิทยาลัยของรัฐ มหาวิทยาลัยเอกชน มหาวิทยาลัยราชภัฏ และมหาวิทยาลัยเทคโนโลยีราชมงคล และ 3. ผลการวิเคราะห์ความถี่ ร้อยละ ของระดับผลการประเมินอภิमान ดังมีรายละเอียดดังต่อไปนี้

1. ผลการประเมินอภิमानของสถาบันอุดมศึกษาในภาพรวม พบว่า มีผลการประเมินอยู่ในระดับดี ($M= 2.78$, $SD=.390$) โดยมาตรฐานที่มีผลการประเมินอยู่ในระดับดีทุกมาตรฐาน คือ มาตรฐานความน่าเชื่อถือของนักประเมิน ($M= 3.21$, $SD=.502$) มาตรฐานจริยธรรมในการประเมิน ($M= 3.11$, $SD=.558$) มาตรฐานอรรถประโยชน์ ($M= 3.08$, $SD=.527$) และมาตรฐานความตรงของการประเมิน ($M= 2.73$, $SD=.573$) ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ ($M= 1.79$, $SD=.526$)

2. ผลการประเมินอภิमानของสถาบันอุดมศึกษา จำแนกตามกลุ่มสถาบัน

2.1 ผลการประเมินอภิमानมหาวิทยาลัยของรัฐ พบว่า มีผลการประเมินอยู่ในระดับดี ($M= 2.80$, $SD=.410$) โดยมาตรฐานความน่าเชื่อถือของนักประเมิน มีผลการประเมินอยู่

ในระดับดีมาก ($M= 3.34$, $SD=.550$) และอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดีมากเช่นกัน ($M= 3.26$, $SD=.441$) และมีผลการประเมินอยู่ในระดับดี คือ มาตรฐานจริยธรรมในการประเมิน ($M= 2.84$, $SD=.833$) และมาตรฐานความตรงของการประเมิน ($M= 2.82$, $SD=.420$) ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับควรปรับปรุง ($M= 1.75$, $SD=.707$)

2.2 ผลการประเมินอภิमानมหาวิทยาลัยเอกชน พบว่า มีผลการประเมินอยู่ในระดับดี ($M= 2.72$, $SD=.409$) โดยมาตรฐานที่มีผลการประเมินอยู่ในระดับดี คือ มาตรฐานความน่าเชื่อถือของนักประเมิน ($M= 3.22$, $SD=.533$) มาตรฐานจริยธรรมในการประเมิน ($M= 3.11$, $SD=.556$) มาตรฐานอรรถประโยชน์ ($M= 2.93$, $SD=.551$) มาตรฐานความตรงของการประเมิน ($M= 2.67$, $SD=.559$) ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับควรปรับปรุง ($M= 1.68$, $SD=.478$)

2.3 ผลการประเมินอภิमानมหาวิทยาลัยราชภัฏ พบว่า มีผลการประเมินอยู่ในระดับดี ($M= 2.79$, $SD=.356$) โดยมาตรฐานที่มีผลการประเมินอยู่ในระดับดี คือ มาตรฐานจริยธรรมในการประเมิน ($M= 3.18$, $SD=.460$) มาตรฐานความน่าเชื่อถือของนักประเมิน ($M= 3.11$, $SD=.471$) มาตรฐานอรรถประโยชน์ ($M= 3.07$, $SD=.522$) และมาตรฐานความตรงของการประเมิน ($M= 2.72$, $SD=.657$) ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ ($M= 1.86$, $SD=.479$)

2.4 ผลการประเมินอภิमानมหาวิทยาลัยเทคโนโลยีราชมงคล พบว่า มีผลการประเมินอยู่ในระดับดี ($M= 2.96$, $SD=.465$) โดยมาตรฐานอรรถประโยชน์ มีผลการประเมินอยู่ในระดับดีมาก ($M= 3.36$, $SD=.513$) และมาตรฐานจริยธรรมในการประเมิน มีผลการประเมินอยู่ในระดับดีมาก ($M= 3.30$, $SD=.322$) มาตรฐานที่มีผลการประเมินอยู่ในระดับดี คือ มาตรฐานความน่าเชื่อถือของนักประเมิน ($M= 3.25$, $SD=.500$) และมาตรฐานความตรงของการประเมิน ($M= 2.89$, $SD=.635$) ส่วนมาตรฐานประสิทธิผลต้นทุนมีผลการประเมินอยู่ในระดับพอใช้ ($M= 2.00$, $SD=.612$)

3. ผลการวิเคราะห์ความถี่ ร้อยละของระดับผลการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันรวมทั้งหมด 50 แห่ง ในภาพรวม พบว่า ส่วนใหญ่อยู่ในระดับดี จำนวน 33 แห่ง (ร้อยละ 66.0) จำแนกเป็น มหาวิทยาลัยราชภัฏ 14 แห่ง มหาวิทยาลัยเอกชน 11 แห่ง มหาวิทยาลัยของรัฐ 5 แห่ง และมหาวิทยาลัยเทคโนโลยีราชมงคล 3 แห่ง ผลการประเมินอยู่ในระดับพอใช้ จำนวน 12 แห่ง (ร้อยละ 24.0) จำแนกเป็นมหาวิทยาลัยเอกชน 5 แห่ง มหาวิทยาลัยราชภัฏ 4 แห่ง มหาวิทยาลัยของรัฐ 2 แห่ง และมหาวิทยาลัยเทคโนโลยีราชมงคล 1 แห่ง และผลการประเมิน

อยู่ในระดับดีมาก จำนวน 5 แห่ง (ร้อยละ 10.0) จำแนกเป็นมหาวิทยาลัยเอกชน 3 แห่ง มหาวิทยาลัยของรัฐและมหาวิทยาลัยเทคโนโลยีราชมงคลกลุ่มละ 1 แห่ง เมื่อพิจารณารายกลุ่มสถาบัน พบว่า มหาวิทยาลัยของรัฐ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 5 แห่ง (ร้อยละ 62.5) รองลงมา อยู่ในระดับพอใช้และดีมาก จำนวน 2 แห่ง และ 1 แห่ง (ร้อยละ 25.0, 12.5) ตามลำดับ มหาวิทยาลัยเอกชน ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 11 แห่ง (ร้อยละ 57.9) รองลงมา อยู่ในระดับพอใช้และดีมาก จำนวน 5 แห่ง และ 3 แห่ง (ร้อยละ 26.3, 15.8) ตามลำดับ มหาวิทยาลัยราชภัฏ ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 14 แห่ง (ร้อยละ 77.8) รองลงมา อยู่ในระดับพอใช้ จำนวน 4 แห่ง (ร้อยละ 22.2) สำหรับมหาวิทยาลัยเทคโนโลยีราชมงคล ส่วนใหญ่มีผลการประเมินอยู่ในระดับดี จำนวน 3 แห่ง (ร้อยละ 60.0) รองลงมา อยู่ในระดับพอใช้และดีมาก จำนวน 1 แห่ง เท่ากัน (ร้อยละ 20.0)

ส่วนที่ 2 ผลตรวจสอบและการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา

ผลการตรวจสอบและสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา แบ่งออกเป็น 2 หัวข้อ คือ 1. ผลการตรวจสอบรายงานผลการประเมินคุณภาพภายใน และ 2. ผลการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ดังมีรายละเอียดดังต่อไปนี้

1. ผลการตรวจสอบรายงานผลการประเมินคุณภาพภายใน

รายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา มีส่วนประกอบใน ส่วนของจุดแข็ง ครบทุกแห่ง (100%) รองลงมาคือ ข้อเสนอแนะ (98%) และแนวปฏิบัติที่ดีหรือแนวทางเสริม (96%) ตามลำดับ ส่วนการตรวจสอบความถูกต้องน่าเชื่อถือของข้อมูลตรวจสอบพบ ในรายงานน้อยที่สุด (46%) เป็นที่น่าสังเกตว่าส่วนสำคัญที่ควรจะมีในรายงานผลการประเมินคุณภาพภายใน ได้แก่ บทนำ และข้อมูลพื้นฐานของสถาบัน ยังพบไม่ครบทุกแห่ง ซึ่งอาจเป็นไปได้ว่า รายงานผลการประเมินคุณภาพภายในของบางแห่งเขียนเป็นบทสรุปสั้นๆ หรือแทรกอยู่ในรายงาน การประเมินตนเอง ทำให้ไม่ใส่รายละเอียดในส่วนนี้

2. ผลการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน

ผลการสังเคราะห์เนื้อหารายงานผลการประเมินคุณภาพภายใน ของสถาบันอุดมศึกษา ซึ่งทำการสังเคราะห์จากรายงานผลการประเมินคุณภาพภายใน และรายงาน การประเมินอภิมานของนักประเมินอภิมาน จำนวน 50 ฉบับ ซึ่งสามารถจำแนกออกเป็น 3 ประเด็น คือ จุดแข็ง จุดบกพร่อง และข้อเสนอแนะ ดังนี้

2.1 จุดแข็ง เนื้อหาของรายงานผลการประเมินคุณภาพภายใน มีจุดแข็งดังนี้ รายงานผลการประเมินคุณภาพภายใน มีการนำเสนอผลการประเมินได้อย่างถูกต้อง เป็นที่ยอมรับ โดยมีการระบุจุดเด่น จุดที่ควรพัฒนาและให้ข้อเสนอแนะที่เป็นประโยชน์ เพื่อให้เห็นทิศทางการวางแผนพัฒนาได้อย่างเหมาะสม เนื่องมาจากการวิเคราะห์และประมวลผลได้อย่างถูกต้องทุกตัวบ่งชี้ การประเมินมีการกำหนดวิธีการประเมินที่สามารถวางแผนและดำเนินการประเมินได้เหมาะสมกับสภาพเป็นจริงและกำหนดช่วงเวลาในการประเมินและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม นอกจากนี้ รายงานการประเมินเขียนด้วยภาษาที่ชัดเจน และมีการนำเสนอด้วยสื่อที่เข้าใจง่าย ซึ่งผู้ประเมินมีความสามารถในการวางแผนและจัดการประเมินอย่างเหมาะสม โดยมีการตรวจสอบความถูกต้องของผลการประเมินจากผู้เกี่ยวข้อง และมีการระบุผู้มีส่วนได้ส่วนเสียในการประเมิน การจัดทำรายงานผลการประเมินเสร็จทันเวลา ซึ่งผู้ประเมินสามารถเข้าใจบริบทและให้ข้อเสนอแนะที่นำไปสู่เป้าหมายที่ต้องการได้และมีจริยธรรม ทำการประเมินรายงานได้นำเชื่อถือ ได้รับการยอมรับจากผู้ที่เกี่ยวข้องในการประเมินได้ อีกทั้งมีการตรวจสอบความน่าเชื่อถือของแหล่งข้อมูล วิธีการเก็บข้อมูล และผู้เก็บรวบรวมข้อมูล รวมถึงมีการเปิดเผยข้อจำกัดที่มีผลกระทบต่อ การประเมิน

2.2 จุดบกพร่อง เนื้อหาของรายงานผลการประเมินคุณภาพภายใน มีจุดบกพร่องดังนี้ รายงานผลการประเมินคุณภาพภายในส่วนใหญ่ พบว่า ไม่ได้ระบุหลักฐานที่แสดงการวางแผนเกี่ยวกับงบประมาณ ไม่มีรายละเอียดการวางแผนจัดการด้านทรัพยากรในการประเมิน การวิเคราะห์ผลการประกันคุณภาพในอดีตซึ่งจะทำให้เกิดการปรับปรุงอย่างต่อเนื่อง ขาดการนำเสนอวัตถุประสงค์ของการประเมิน และไม่ได้แสดงเอกสารหลักฐานที่ตกลงทำการประเมินระหว่างสถาบัน นอกจากนี้ยังพบว่ารายงานผลการประเมินบางส่วน รูปแบบของรายงานการประเมินคุณภาพภายในของคณะกรรมการ ฯ ไม่มีการแสดงถึงรายละเอียดชัดเจน ไม่ครบถ้วน ไม่แสดงการนำเสนอร่องรอย /หลักฐาน/ รายละเอียดการสัมภาษณ์ และการบันทึกข้อมูลภาคสนามที่พบและนำมาใช้ประกอบการพิจารณาตัดสินผลการประเมินแต่ละตัวบ่งชี้ ไม่ปรากฏหลักฐานที่เกี่ยวข้องกับการวางแผนการประเมินและผู้มีส่วนได้ส่วนเสียไว้ และยังขาดการตรวจสอบความถูกต้องของข้อมูล ขาดข้อมูลพื้นฐาน ภูมิหลังของสถาบัน รวมถึงไม่มีการระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่ชัดเจนและเหมาะสม

2.3 ข้อเสนอแนะ

ข้อเสนอแนะจากการประเมินคุณภาพภายใน ระดับอุดมศึกษา มีรายละเอียดดังต่อไปนี้ การประเมินคุณภาพภายใน และการรายงานผลการประเมินคุณภาพภายใน ควรปรับปรุงพัฒนารูปแบบรายงานการประเมินที่ครอบคลุม ครบถ้วน มีสารสนเทศที่เพียงพอต่อการนำเสนอให้สถาบันอุดมศึกษาและผู้เกี่ยวข้องตลอดจนสาธารณชนได้เข้าใจมากขึ้น ข้อเสนอหรือจุดที่ควรพัฒนาควรระบุสาเหตุของปัญหาและข้อเสนอแนะอย่างชัดเจน ควรปรับปรุงคุณภาพของผู้ประเมินให้มีความรู้ ความสามารถ และทักษะในการประเมินที่ถูกต้องและได้มาตรฐานที่ใกล้เคียง เพื่อประโยชน์ของคุณภาพการศึกษาต่อไป การประเมินในครั้งนี้อาจประเมินได้โดยตรงจากผู้ประเมิน แต่ดูจากเอกสารรายงานอย่างเดียว ทำให้อาจประเมินได้ไม่ชัดเจนตามความเป็นจริง และควรมีการเปิดเผยรายงานผลการประเมินให้ครบถ้วน และเผยแพร่เอกสารการประเมินให้สาธารณชนรับทราบ ทั้งนี้ควรจัดหาผู้ประเมินที่ไม่มีอิทธิพลเกี่ยวกับผลประโยชน์จากกลุ่มใดที่ส่งผลกระทบต่อประเมิน

อภิปรายผลการวิจัย

การอภิปรายผลการวิจัยในครั้งนี้ ผู้วิจัยแยกอภิปรายผลตามวัตถุประสงค์การวิจัย 3 ข้อ คือ 1) การพัฒนามาตรฐานการประเมินอภิมาน สำหรับประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษาในประเทศไทย 2) การตรวจสอบคุณภาพมาตรฐานการประเมินอภิมาน 3) ผลจากการประยุกต์ใช้มาตรฐานการประเมินอภิมานในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา ดังมีรายละเอียดต่อไปนี้

1. การพัฒนามาตรฐานการประเมินอภิมาน สำหรับประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษาในประเทศไทย

มาตรฐานการประเมินอภิมาน ที่พัฒนาขึ้นได้ทำการศึกษาจากเอกสารและงานวิจัยที่เกี่ยวข้องกับการพัฒนามาตรฐานการประเมิน จากหลากหลายแนวคิด ตั้งแต่ มาตรฐานการประเมิน Program Evaluation Standards ตามแนวทางของ the Joint Committee on Standard for Educational Evaluation (1994, และ 2010) ประกอบด้วย 4 มาตรฐานหลัก 30 มาตรฐานย่อย ที่มีหลายๆ ประเทศนำไปประยุกต์ใช้เป็นมาตรฐานการประเมินในบริบทของตนเอง ได้แก่ 1) The African Evaluation guidelines 2002 (AEG) ที่พัฒนาขึ้นโดยสมาคมการประเมินของประเทศแอฟริกา (African Evaluation Association) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 30 มาตรฐานย่อย 2) The DeGEval's Standards (2003) ที่พัฒนาโดยสมาคมการประเมินของประเทศเยอรมัน (German Evaluation Society) กำหนดมาตรฐานหลัก 4 มาตรฐาน และ 25 มาตรฐานย่อย และ 3)

SEVAL's Evaluation Standards พัฒนาโดยสมาคมการประเมินของประเทศสวิสเซอร์แลนด์ (Swiss Evaluation Society) กำหนดมาตรฐานการประเมินหลัก 4 มาตรฐาน และ 28 มาตรฐานย่อย (Laubli Loud, 2004) นอกจากนี้ในส่วนของสมรรถนะของผู้ประเมิน ได้ศึกษาจากแนวคิดของมาตรฐานของผู้ประเมิน ได้แก่ 1) Guiding Principles for Evaluators ของ สมาคมประเมินแห่งสหรัฐอเมริกา (American Evaluation Association [AEA] ,2003) 2) ทักษะที่จำเป็นในการประเมิน Essential Skills Series in Evaluation ของสมาคมประเมินแห่งแคนาดา (Canadian Evaluation Society, 1999) และ 3) สมรรถนะที่จำเป็นสำหรับผู้ประเมินโครงการ Essential Competencies for Program Evaluators (Ghere & et al., 2006)

ทั้งนี้ได้ใช้กรอบแนวทางในการพัฒนามาตรฐานการประเมินอภิมาน ตามมาตรฐานการประเมินอภิมาน 5 มาตรฐานของ Davidson (2005) คือ 1) ความตรงของการประเมิน (validity) 2) อรรถประโยชน์ (utility) 3) จริยธรรมในการประเมิน (ethicality) 4) ความน่าเชื่อถือของนักประเมิน (credibility) และ 5) ประสิทธิภาพต้นทุน (cost-effectiveness) นอกจากนี้มาตรฐานการประเมินอภิมาน 5 มาตรฐาน ดังกล่าวได้มีการนำไปใช้เป็นกรอบในการประเมินทุนวิจัยของรัฐบาล เปรียบเทียบระหว่างประเทศใน 16 ประเทศ เพื่อให้พิจารณาความเหมาะสมของโมเดลการประเมินอีกด้วย (Coryn & et.al, 2007; Coryn & Scriven, 2007) ดังนั้น จึงสังเคราะห์ได้กรอบมาตรฐานการประเมินอภิมานที่มีความสมบูรณ์ ประกอบด้วย 38 ตัวบ่งชี้ โดยแนวคิดดังกล่าวได้นำเสนอที่ประชุม The American Evaluation Association (AEA) จัดขึ้นที่ The world class Rosen Shingle Creek Resort, Orlando, Florida ประเทศสหรัฐอเมริกา ระหว่างวันที่ 11 - 14 พฤศจิกายน พ.ศ. 2552 นักวิชาการให้ความเห็นว่าสามารถนำแนวคิดดังกล่าวไปประยุกต์ใช้ในการประเมินในระดับสากล หรือประเทศอื่น ๆ ได้

สำหรับการพัฒนาตัวบ่งชี้ในการประเมินอภิมาน เพื่อใช้ในการประเมินรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษาในประเทศไทย ประกอบด้วย 38 ตัวบ่งชี้ ซึ่งได้ศึกษาสังเคราะห์จาก มาตรฐานการประเมิน ดังที่กล่าวมาแล้วใน 7 แนวทาง และนอกจากนี้ยังพิจารณาการพัฒนาเกณฑ์การประเมินจากมาตรฐานและตัวบ่งชี้ในการประเมินอภิมานรายงานการประเมินคุณภาพภายในระดับอุดมศึกษา ของสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) (2551) และคู่มือการประกันคุณภาพภายในสถานศึกษาระดับอุดมศึกษา ของสำนักงานคณะกรรมการการอุดมศึกษา (2552) พบว่า มาตรฐานการประเมินอภิมานที่พัฒนาขึ้นมีจุดเด่นที่พิเศษกว่า มาตรฐานการประเมินของ program evaluation standard ก็คือ สามารถประเมินได้ครอบคลุมถึงมาตรฐานการประเมินของผู้ประเมิน Guiding Principles for Evaluators ของ AEA

Essential Skills Series in Evaluation ของแคนาดา และ Essential Competencies for Program Evaluators จากงานวิจัยของ Ghere & et al. (2006) การประเมินอภิมานของ สมศ. ระดับอุดมศึกษา และนอกจากนี้ ตัวอย่างที่ Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา ยังเหมาะสมกับการประเมินในบริบทไทยที่คำนึงถึงกฎกระทรวง การรายงานผลการประกันคุณภาพการศึกษา พ.ศ. 2546 ของสำนักงานคณะกรรมการการอุดมศึกษา กำหนดไว้ว่าสถาบันต้องจัดส่งเอกสารรายงานการประเมินให้เสร็จสิ้นภายใน 120 วัน หลังสิ้นปีการศึกษา สามารถใช้ในการกำกับ ติดตาม ประเมิน รายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษาได้อีกด้วย

แต่ในส่วนของมาตรฐานประสิทธิผลต้นทุน (cost-effectiveness) พบว่า มีผลการประเมินต่ำในทุกกลุ่มสถาบัน ทั้งนี้เนื่องมาจากยังเป็นส่วนที่ไม่มีข้อกำหนดไว้ในกรเขียนรายงานประเมินคุณภาพภายใน และส่วนมาก สถาบันไม่มีการระบุรายละเอียดไว้ในรายงาน จึงทำให้ไม่สามารถประเมินได้โดยตรงจากรายงาน และไม่สามารถบอกได้ว่า การประเมินคุณภาพภายในที่จัดขึ้นตามกรอบแนวทางที่ต้นสังกัดกำหนดนั้น มีความคุ้มค่าต่อการจัดการประเมินหรือไม่ อย่างไร แต่ทั้งนี้ที่นำมาพิจารณาประเมินอภิมานเนื่องจากมีข้อกำหนดไว้ในมาตรฐานการประเมิน program evaluation standards และมาตรฐานของผู้ประเมิน Essential Skills Series in Evaluation (Canadian Evaluation Society, 1999) และ Essential Competencies for Program Evaluators (Ghere & et al., 2006) มาตรฐานประสิทธิผลต้นทุน จึงมีความสำคัญ ควรนำไปประกอบการพิจารณาประเมินด้วย

การพิจารณาค่าน้ำหนักองค์ประกอบของมาตรฐานการประเมินอภิมาน พบว่า เมื่อทำการวิเคราะห์องค์ประกอบเชิงยืนยันจากข้อมูลเชิงประจักษ์ พบว่า ค่าน้ำหนักในแต่ละมาตรฐานมีค่า ดังนี้ มาตรฐานอรรถประโยชน์ (utility) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.90 รองลงมาคือ มาตรฐานความตรงของการประเมิน (validity) ความน่าเชื่อถือของนักประเมิน (credibility) และ จริยธรรมในการประเมิน (ethicality) มีค่าน้ำหนักองค์ประกอบเท่ากับ 0.86, 0.82 และ 0.40 ตามลำดับ และประสิทธิผลต้นทุน (cost-effectiveness) มีค่าน้ำหนักน้อยที่สุดเท่ากับ 0.17 ซึ่งพบว่า น้ำหนักในมาตรฐานความตรงของการประเมินและอรรถประโยชน์ มีความสอดคล้องกับที่งานวิจัยกำหนดไว้ แต่มาตรฐานจริยธรรมในการประเมิน และความน่าเชื่อถือของนักประเมินมีความแตกต่างกัน กล่าวคือ งานวิจัยกำหนดไว้ ร้อยละ 24 และ 8 ตามลำดับ ทั้งนี้เมื่อนำค่าน้ำหนักองค์ประกอบตามที่วิเคราะห์ที่ได้มาคิดเป็นร้อยละ ตามค่าน้ำหนักที่คำนวณได้ จึงควรกำหนดค่าน้ำหนักในแต่ละมาตรฐานใหม่ ดังต่อไปนี้คือ มาตรฐานความตรงของการประเมินร้อยละ 30

อรรถประโยชน์ร้อยละ 30 จริยธรรมในการประเมินร้อยละ 10 ความน่าเชื่อถือของนักประเมินร้อยละ 25 ส่วนประสิทธิผลต้นทุน ร้อยละ 5

2. การตรวจสอบคุณภาพของมาตรฐานการประเมินอภิมาน

มาตรฐานการประเมินอภิมาน ที่พัฒนาขึ้น ได้นำไปพัฒนาตัวบ่งชี้ และเกณฑ์การประเมินอภิมาน โดยจัดทำเครื่องมือ คือ แบบประเมินอภิมาน และคู่มือการประเมินอภิมาน ซึ่งได้ผ่านกระบวนการพัฒนาหลายขั้นตอน และมีการพัฒนาปรับปรุงอยู่ตลอดเวลา ทั้งนี้เครื่องมือในการประเมินอภิมาน ได้ผ่านการพิจารณาจากผู้เชี่ยวชาญด้านการประกันคุณภาพ และด้านการวัดผลในเบื้องต้น ต่อจากจึงนำไปทดลองใช้ในการประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษาจากผู้เข้าร่วมอบรมในหลักสูตรพัฒนานักประเมินอภิมาน เพื่อพิจารณาถึงปัญหา อุปสรรคในการใช้เครื่องมือประเมินอภิมาน นอกจากนี้ได้ทำการวิเคราะห์หาคุณภาพของเครื่องมือประเมินอภิมาน ในด้านความตรง (validity) และความเที่ยง (reliability) ดังต่อไปนี้

2.1 คุณภาพด้านความตรง (validity) เมื่อนำผลการประเมินอภิมานที่ประเมินโดยนักประเมินอภิมาน จำนวน 10 ฉบับ ไปเปรียบเทียบกับผลการบ่งชี้คุณภาพของรายงานผลการประเมินคุณภาพภายในระดับอุดมศึกษา ฉบับเดียวกัน โดยผู้ทรงคุณวุฒิ ด้วยเทคนิคกลุ่มรู้ชัด (known group technique) พบว่า สามารถจำแนกความแตกต่างของคุณภาพรายงานผลการประเมินคุณภาพภายในออกเป็นรายงานที่มีคุณภาพดี และรายงานที่มีคุณภาพต่ำได้ โดยสอดคล้องกับคุณภาพของรายงานที่ผู้ทรงคุณวุฒิบ่งชี้คุณภาพไว้ แสดงว่า คุณภาพของเครื่องมือประเมินอภิมานที่พัฒนาขึ้นตามกรอบมาตรฐานการประเมินอภิมาน มีคุณภาพด้านความตรงเชิงโครงสร้าง กับการประเมินคุณภาพโดยผู้ทรงคุณวุฒิด้านการประกันคุณภาพ ระดับอุดมศึกษา ที่มีความเชี่ยวชาญในการประเมินคุณภาพภายใน ระดับสถาบัน ซึ่งสอดคล้องกับงานวิจัยของอภิสิทธิ์ ทรงบัณฑิตย์ (2550) พบว่า ผลการเปรียบเทียบผลการประเมินของสถานศึกษาตามเทคนิคกลุ่มรู้ชัด (known group technique) โดยเทียบจากค่าคะแนนประเมินตามระบบมัลคอม บัลดริจ สถานศึกษาที่มีลักษณะการวัดเด่นชัด มีคะแนนรวมตามระบบมัลคอม บัลดริจแตกต่างจากสถานศึกษาเปรียบเทียบ

2.2 คุณภาพด้านความเที่ยง (reliability) ด้วยการวิเคราะห์ค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) จากข้อมูลการประเมินรายงานผลการประเมินคุณภาพภายใน จำนวน 5 ฉบับ กรณีใช้นักประเมินอภิมานประเมิน 2 คน ประเมินรายงานฉบับเดียวกัน โดยการวิธีการหาความเที่ยงโดยวิธีสัมประสิทธิ์การสรุปอ้างอิง (G-Theory) I X (R:P) Design เมื่อทำการประเมินรายงานผลการประเมินคุณภาพภายในทั้ง 5 ฉบับ โดยใช้ผู้ประเมินอภิมานต่างกัน

จำนวน 2 คน ด้วยการให้แบบประเมินอภิมานเหมือนกันจำนวน 38 ตัวบ่งชี้ กรณี เพิ่มจำนวนนักประเมินอภิมาน และเพิ่มจำนวนรายงานผลการประเมินคุณภาพภายใน มีค่าสัมประสิทธิ์การสรุปอ้างอิง (generalizability coefficient) มีค่ามากขึ้นใน 3 แนวทาง ดังนี้ 1) ควรกำหนดนักประเมินอภิมานจำนวน 3 คนขึ้นไป แตกต่างกัน ประเมินรายงานผลการประเมินฯฉบับเดียวกันไขว้กัน จำนวน 3 ฉบับ 2) ควรกำหนดนักประเมินอภิมานจำนวน 2 คนขึ้นไป แตกต่างกัน ประเมินรายงานผลการประเมินฯฉบับเดียวกัน จำนวน 5 ฉบับ และ 3) ควรใช้นักประเมินอภิมานตั้งแต่ 1 คนขึ้นไป ประเมินรายงานฯแต่ละฉบับ ไม่ซ้ำกัน จำนวน 7 และ 9 ฉบับ ตามลำดับ เนื่องจากมีค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสัมบูรณ์ มากกว่า 0.70 สอดคล้องกับงานวิจัยของ Lei, Smith, และ Suen (2007) ที่ทำการศึกษาการประมาณค่าสัมประสิทธิ์การสรุปอ้างอิง กรณีใช้ผู้สังเกตในการวิจัยพบว่า เมื่อใช้ผู้สังเกต 1 คน มีค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสัมพัทธ์ และเชิงสัมบูรณ์ เท่ากับ 0.68 และ 0.66 ตามลำดับ แต่เมื่อเพิ่มผู้สังเกตเป็น 2 คน ค่าสัมประสิทธิ์การสรุปอ้างอิงเชิงสัมพัทธ์ และเชิงสัมบูรณ์ เพิ่มขึ้น เป็น 0.81 และ 0.79 ตามลำดับ

3. ผลจากการประยุกต์ใช้มาตรฐานการประเมินอภิมานในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา

การอภิปรายผลในส่วนของการประยุกต์ใช้มาตรฐานการประเมินอภิมานในการประเมินคุณภาพรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา แบ่งออกเป็น 2 ประเด็น คือ การอภิปรายผลการวิเคราะห์เชิงปริมาณ และการอภิปรายผลวิเคราะห์เชิงเนื้อหา ดังมีรายละเอียดต่อไปนี้

3.1 การอภิปรายผลวิเคราะห์เชิงปริมาณ

ผลการประยุกต์ใช้มาตรฐานการประเมินอภิมาน โดยการประเมินอภิมานรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษาอยู่ในระดับดี ทั้งในภาพรวมและระดับกลุ่มสถาบัน โดยที่ ในภาพรวมมาตรฐานด้านความน่าเชื่อถือของนักประเมินมีค่าสูงสุด เนื่องจากผู้ประเมินที่ทำหน้าที่เป็นผู้ประเมินในระดับสถาบัน มาจากการอบรมหลักสูตรอบรมผู้ประเมินคุณภาพภายใน ของสำนักงานคณะกรรมการการอุดมศึกษา และนอกจากนี้ผู้ประเมินคุณภาพภายใน ระดับสถาบัน ที่อย่างน้อยร้อยละ 50 มาจากภายนอกสถาบัน เพื่อเป็นการควบคุมคุณภาพการประเมินคุณภาพภายใน (สำนักงานคณะกรรมการการอุดมศึกษา, 2552) โดยเป็นผู้ที่มีความเชี่ยวชาญเฉพาะด้านที่สอดคล้องกับการจัดการศึกษาของสถาบันอุดมศึกษานั้น ทำให้เห็นมุมมองในการบริหารจัดการ การเรียนการสอนของสถาบันได้เด่นชัดมากขึ้น เมื่อพิจารณารายกลุ่มสถาบันพบว่า มาตรฐานด้านความ

น่าเชื่อถือของนักประเมินมีค่าสูงสุดเช่นกัน ยกเว้น มหาวิทยาลัยเทคโนโลยีราชมงคลที่มาตรฐานด้านอรรถประโยชน์สูงสุด อยู่ในระดับดีมาก ทั้งนี้เนื่องมาจาก มีการระบุช่วงเวลา ระยะเวลาในการประเมิน และการเผยแพร่รายงานผลการประเมิน อยู่ในระดับดีมาก ซึ่งการเผยแพร่ผลการประเมินได้ทันเวลาที่กำหนด การเขียนรายงานมีความชัดเจน สามารถเข้าใจได้ทุกกลุ่ม เนื่องจากมหาวิทยาลัยเทคโนโลยีราชมงคล มีรูปแบบการบริหารจัดการที่เป็นแนวทางเดียวกัน มีระบบการประกันคุณภาพ การเขียนรายงานร่วมกัน มีการเผยแพร่ผลการประเมินอย่างเป็นระบบ เข้าใจง่ายสืบค้นได้ง่าย เพราะว่ามีมหาวิทยาลัยเทคโนโลยีราชมงคล มีการแยกตัวออกเป็นมหาวิทยาลัย 9 แห่ง ในระยะเวลาไม่นาน ดังนั้นจึงยังมีส่วนดีที่มีการจัดการร่วมกันอยู่ จึงค่อนข้างมีความเป็นเอกภาพมากกว่ากลุ่มอื่น

สำหรับมาตรฐานด้านความตรงในการประเมิน ด้านอรรถประโยชน์ และด้านจริยธรรมในการประเมินอยู่ในระดับดีขึ้นไปทุกกลุ่มสถาบัน ซึ่งพบว่า มาตรฐานด้านความตรงในการประเมิน มีการรายงานผลการประเมินจากข้อค้นพบทั้งเชิงบวกและเชิงลบ มีผลการประเมินสูงสุด เนื่องจากนักประเมินอภิมาน สามารถนำเสนอทั้งจุดเด่น จุดที่ควรพัฒนาและข้อจำกัดในการประเมิน ส่วนคุณภาพของสารสนเทศมีผลการประเมินต่ำสุด ทั้งนี้ อาจเนื่องมาจากการประเมินคุณภาพภายใน ระดับสถาบัน สำนักงานคณะกรรมการการอุดมศึกษา เป็นผู้กำหนดกรอบ แนวทาง มาตรฐาน เกณฑ์ในการประเมินคุณภาพภายในไว้แล้ว รวมทั้งเครื่องมือที่ใช้ในการประเมิน ดังนั้นผู้ประเมินจึงใช้เครื่องมือที่กำหนดให้เหมือนกัน นอกจากนี้ในการประเมินอภิมานไม่เห็นกระบวนการในการตรวจสอบคุณภาพของสารสนเทศ เพราะผู้ประเมินไม่ได้ระบุรายละเอียดกลุ่มที่ใช้ และกระบวนการในการตรวจสอบไว้ในรายงาน สำหรับมาตรฐานด้านอรรถประโยชน์ในการประเมิน มีการเผยแพร่ผลการประเมินสูงสุดสอดคล้องกับภาพรวม และการกำหนดเทคนิคในการเก็บรวบรวมข้อมูลต่ำสุด ทั้งนี้รายงานฯ ส่วนใหญ่ไม่มีระบุรายละเอียดไว้ในการวางแผนการประเมินที่ชัดเจน ที่สามารถแสดงให้เห็นว่า ผู้ประเมินวางแผนกำหนดเทคนิคในการรวบรวมสารสนเทศที่จำเป็น ตามข้อจำกัดของทรัพยากร ส่วนมาตรฐานด้านจริยธรรมในการประเมิน มีผลการประเมินเป็นที่ยอมรับมีค่าสูงสุด เนื่องจากผลการประเมินไม่ได้รับอิทธิพลทางการเมืองที่ขัดขวางการประเมิน รวมถึงผลการประเมินสะท้อนให้เห็นจุดเด่น/จุดที่ควรพัฒนา/ทิศทางการพัฒนาในอนาคตได้ การติดต่อกับผู้ที่เกี่ยวข้อง โดยคำนึงถึงความแตกต่าง หลากหลายของผู้ที่เกี่ยวข้องต่ำสุด เนื่องจากไม่หลักฐานหรือการอธิบายถึงการศึกษาถึงภูมิหลังทางสังคมและวัฒนธรรมของผู้ที่เกี่ยวข้องกับการประเมิน รายงานไม่ได้ระบุไว้จึงทำให้ไม่สามารถรู้ได้ว่ามีการศึกษาภูมิหลังของผู้ที่เกี่ยวข้อง

ส่วนมาตรฐานด้านประสิทธิผลต้นทุนในภาพรวมอยู่ในระดับพอใช้ เพียงมาตรฐานเดียว ทั้งนี้ตามที่กล่าวไปแล้วนั้นว่า ประสิทธิภาพต้นทุนไม่ได้ระบุไว้ในรูปแบบการเขียน

รายงานของสำนักงานคณะกรรมการการอุดมศึกษาว่า ให้มีการเขียนรายงานเกี่ยวกับ ต้นทุนงบประมาณ และทรัพยากรที่ใช้ในการประเมินคุณภาพภายใน ซึ่งในความเป็นจริงแล้ว การประเมินคุณภาพภายใน แต่ละปีจำเป็นต้องมีการวางแผนการใช้งบประมาณทั้ง ด้านการเงิน บุคลากร วัสดุอุปกรณ์และสถานที่ จำนวนมากพอสมควร เพื่อให้ได้ผลการประเมินคุณภาพภายใน ที่มีคุณภาพและสามารถนำผลการประเมินมาใช้ในการพัฒนาปรับปรุงคุณภาพของสถาบันอุดมศึกษาได้ ซึ่งจากการสัมภาษณ์ผู้ทรงคุณวุฒิด้านการประกันคุณภาพ ได้ให้ความเห็นว่า สถาบันอุดมศึกษาสามารถที่จะเลือกสรรผู้ประเมินคุณภาพภายใน ที่มีความเชี่ยวชาญในการประเมินคุณภาพภายใน และสามารถให้ข้อเสนอแนะที่เป็นประโยชน์ต่อสถาบัน ทั้งมีอาจจำเป็นต้องมีต้นทุนการประเมินสูง เพื่อให้ได้ผลการประเมินที่มีคุณภาพ และทั้งนี้ควรมีการหมุนเวียน สับเปลี่ยนผู้ประเมินคุณภาพภายใน ที่มาจากนอกสถาบันเพื่อให้มีมุมมองในการประเมินที่หลากหลาย

3.2 การอภิปรายผลวิเคราะห์เชิงเนื้อหา

การวิเคราะห์ข้อมูลเชิงเนื้อหา จากการตรวจสอบรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา พบว่า มีรายงานส่วนใหญ่ยังมีส่วนประกอบไม่ปฏิบัติตามรูปแบบที่สำนักงานคณะกรรมการการอุดมศึกษากำหนดไว้ ประเด็นที่สำคัญคือ ไม่มีกระบวนการตรวจสอบความถูกต้องของข้อมูลนำเชื่อถือของข้อมูล ทำให้ไม่สามารถทราบได้ว่าข้อมูลที่ใช้ในการตัดสินผลการประเมินเพียงพอ และน่าเชื่อถือ และรายงานการประเมินที่ดีนั้น ควรระบุบทเรียนที่ได้รับจากการประเมินในครั้งนี้ด้วย เพื่อให้ข้อเสนอแนะที่เป็นประโยชน์ต่อการประเมินในครั้งต่อไป หรือการปรับปรุงพัฒนางานประเมินให้มีคุณภาพมากยิ่งขึ้น อีกทั้งควรมีการตรวจสอบจากผู้ที่เกี่ยวข้องกับหลายฝ่ายอีกด้วย ตามที่ UNDP (United Nations Development Programme) (2008), USAID (United States Agency International Development) (2006) และ Standard format for evaluation reports (Lisa Emerson, 1995) กำหนดไว้ รวมทั้งงานวิจัยของ Lin, Lin, และ Chern. (2009) ที่ได้กำหนดให้มีการทำรายงานการประเมินผลิตภัณฑ์ฯ เพื่อพัฒนาคุณภาพของยาให้เป็นที่ยอมรับ

การประเมินคุณภาพภายใน ระดับอุดมศึกษา มีการกำหนดกรอบแนวทาง และช่วงเวลาในการประเมิน และแนวทางการกำหนดคณะผู้ประเมินคุณภาพภายใน โดยสำนักงานคณะกรรมการการอุดมศึกษา เพื่อให้สถาบันอุดมศึกษาดำเนินการตามที่กำหนด อีกทั้งมีการอบรมพัฒนาผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษาที่มีความรู้ ความสามารถในมาตรฐาน ตัวบ่งชี้ และเกณฑ์การประเมินคุณภาพภายใน ทำให้เป็นจุดแข็งของการดำเนินการประเมินคุณภาพ

ภายใน จึงทำให้รายงานผลการประเมินส่วนใหญ่ มีการนำเสนอผลการประเมินได้อย่างถูกต้อง เป็นที่ยอมรับ โดยมีการระบุจุดเด่น จุดที่ควรพัฒนาและให้ข้อเสนอแนะที่เป็นประโยชน์ วิธีการประเมินที่สามารถวางแผนและดำเนินการประเมินได้เหมาะสมกับสภาพเป็นจริงและกำหนดช่วงเวลาในการประเมินและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม โดยมีการกำกับติดตามจากหน่วยงานต้นสังกัดอยู่แล้ว แต่ทั้งนี้ประเด็นสำคัญที่สำนักงานคณะกรรมการการอุดมศึกษา ไม่ได้กำหนดให้เสนอรายละเอียดเกี่ยวกับวางแผนเกี่ยวกับงบประมาณ ซึ่งสถาบันอุดมศึกษาต่างๆ ต้องมีการดำเนินการจัดทำแผนงบประมาณสำหรับการประเมินคุณภาพภายใน อยู่แล้ว แต่ไม่ได้มีระบุไว้ในรายงาน จึงไม่พบรายละเอียด ทำให้ไม่สามารถประเมินได้ว่างานประเมินมีประสิทธิภาพเพียงพอหรือไม่ เหมาะสมสอดคล้องกับงบประมาณและทรัพยากรที่ใช้ไปในการประเมินคุณภาพภายใน รวมทั้งข้อเสนอแนะ และประโยชน์ที่ได้จากการประเมินมีคุณค่า และคุ้มค่ากับค่าใช้จ่าย ทั้งเรื่องเงิน เวลา บุคลากร และสถานที่ที่ใช้ไปหรือไม่ ซึ่งสอดคล้องกับผลการวิเคราะห์ในเชิงปริมาณ ที่พบว่า มาตรฐานด้านประสิทธิภาพในการประเมินมีระดับควรปรับปรุงถึงพอใช้ นอกจากนี้ประเด็นของการวิเคราะห์ผลการประกันคุณภาพในอดีต ซึ่งจะทำให้เกิดการปรับปรุงอย่างต่อเนื่อง ซึ่งพบว่างานประเมินคุณภาพภายใน ส่วนใหญ่ไม่ได้กล่าวถึงการนำผลการประเมินที่ผ่านมาไปใช้ประโยชน์ในการวางแผน พัฒนาปรับปรุงสถาบัน ทำให้งานประเมินเกิดประโยชน์น้อยมาก และได้มีการพิจารณาอย่างจริงจังในการนำข้อเสนอแนะไปพิจารณาปรับปรุง หรือวางแผนการพัฒนาอย่างไร อีกทั้งข้อเสนอแนะบางประเด็นของคณะกรรมการเป็นสิ่งเป็นนามธรรม ปฏิบัติได้ยาก ไม่สอดคล้องกับบริบทของสถาบัน ขึ้นอยู่กับทักษะและความเชี่ยวชาญของคณะผู้ประเมินคุณภาพภายใน ดังนั้นควรพิจารณาถึงประเด็นของรายงานการประเมินในรายละเอียดที่ควรระบุ และเชื่อมโยงให้เห็นในภาพรวม ความเกี่ยวข้องของข้อมูลที่แสดงให้เห็นถึงอัตลักษณ์ของสถาบันในภาพรวม เพื่อเป้าหมายการพัฒนาสถาบันให้บรรลุเป้าหมายสูงสุดได้ โดยเฉพาะสถาบันที่เปิดใหม่ และสถาบันที่เกิดจากการรวมกลุ่มของวิทยาลัยต่างๆ จัดตั้งขึ้นเป็นมหาวิทยาลัยให้เกิดความเป็นเอกภาพอย่างชัดเจน

นอกจากนี้ควรพิจารณาทบทวน ปรับปรุงแนวทางการเขียนรายงานการประเมินคุณภาพภายใน ที่มีรายละเอียดที่ครอบคลุม ครบถ้วน มีสารสนเทศที่เพียงพอเหมาะสมกับลักษณะของสถาบันนั้นๆ อีกทั้งควรมีการอบรมทบทวนผู้ประเมินคุณภาพภายใน เพื่อปรับทิศทางการประเมินให้มีแนวทางเดียวกัน เข้าใจตรงกันและมีมาตรฐานการตัดสินที่สอดคล้องกัน รวมทั้งมีการตรวจสอบซ้ำผลการประเมินคุณภาพภายใน จากหลายๆฝ่ายที่เกี่ยวข้องก่อนที่มีการเผยแพร่

ผลการประเมินผู้สําคาารณชน เพื่อประโยชน์ของสถาบันอุดมศึกษาเอง และประโยชน์ของผู้มีส่วนได้ส่วนเสียทุกกลุ่มที่เกี่ยวข้อง

ข้อเสนอแนะ

ข้อเสนอแนะในการวิจัย ผู้วิจัยขอเสนอแนะใน 3 ประเด็น คือ 1) ข้อเสนอแนะเชิงนโยบาย 2) ข้อเสนอแนะในการนำผลการวิจัยไปใช้ประโยชน์ และ3) ข้อเสนอแนะสำหรับการวิจัยในครั้งต่อไป ดังมีรายละเอียดต่อไปนี้

ข้อเสนอแนะเชิงนโยบาย

1. หน่วยงานต้นสังกัด สำนักงานคณะกรรมการการอุดมศึกษาและสถาบันอุดมศึกษา ควรส่งเสริมให้มีการตรวจสอบผลการประเมิน หรือส่งเสริมให้มีการประเมินอภิมานภายในหน่วยงาน เพื่อปรับปรุงคุณภาพการประเมิน ผู้ประเมิน และการเขียนรายงานการประเมินให้มีคุณภาพมากขึ้น

2. สถาบันอุดมศึกษา ควรมีนโยบายในการวางแผนงบประมาณ และทรัพยากรที่จำเป็นเพื่อการประกันคุณภาพภายใน และนำเสนอรายละเอียดเกี่ยวกับงบประมาณ ค่าใช้จ่ายที่ใช้ในการประกันคุณภาพภายใน เพื่อสามารถประเมินถึงประสิทธิภาพของต้นทุนในการประกันคุณภาพภายใน ได้อย่างเหมาะสม

3. หน่วยงานต้นสังกัด ควรกำหนดนโยบายให้สถาบันอุดมศึกษา มีการจัดทำแผนการดำเนินงาน และแผนกลยุทธ์ที่มาจากข้อเสนอแนะในการประกันคุณภาพภายใน เพื่อส่งเสริมผลักดัน ให้นำข้อเสนอแนะจากผู้ประเมินไปใช้ประโยชน์ และเป็นการส่งเสริมให้ผู้ประเมินคุณภาพภายใน พิจารณาให้ข้อเสนอแนะที่สามารถนำไปใช้ในการวางแผนพัฒนาสถาบันได้ อย่างเป็นรูปธรรม

ข้อเสนอแนะในการนำผลการวิจัยไปใช้ประโยชน์

1. หน่วยงานต้นสังกัด สำนักงานคณะกรรมการการอุดมศึกษา และสถาบันอุดมศึกษา ควรมีการพิจารณากรอบแนวทาง และรูปแบบการเขียนรายงานการประเมินคุณภาพภายใน ให้มีความชัดเจน ครอบคลุมมาตรฐานการประเมิน โดยคำนึงรายละเอียดรูปแบบมาตรฐานการเขียนรายงานการประเมินในระดับสากลมาประกอบการพิจารณา ทั้งนี้ เพื่อสามารถให้สารสนเทศที่เป็นประโยชน์สอดคล้องกับการพัฒนาคุณภาพของสถาบันอุดมศึกษา รวมทั้งทบทวนความเข้าใจของผู้ประเมินคุณภาพภายใน ที่ผ่านการอบรมหลักสูตรผู้ประเมินคุณภาพภายใน ให้

ตรงกัน ในเรื่องของการเขียนรายงานการประเมินที่เป็นประโยชน์ต่อการพัฒนาสถาบันอุดมศึกษา รวมทั้งให้ข้อเสนอแนะที่สามารถนำไปปฏิบัติได้จริง เหมาะสม สอดคล้องกับบริบทและอัตลักษณ์ของสถาบันอุดมศึกษานั้น

2. สถาบันอุดมศึกษาในแต่ละกลุ่ม มหาวิทยาลัยของรัฐ มหาวิทยาลัยเอกชน มหาวิทยาลัยราชภัฏ และมหาวิทยาลัยเทคโนโลยีราชมงคล ควรนำผลการประเมินอภิมานไปพิจารณาปรับปรุงคุณภาพการประเมินคุณภาพภายใน ในทุกมาตรฐาน เพื่อให้การประเมินคุณภาพภายใน ระดับสถาบัน สามารถประเมินได้ครอบคลุม และพัฒนารายงานการประเมินตนเองให้เป็นไปตามมาตรฐานการประเมิน เพื่อให้มีข้อมูลครบถ้วน เพียงพอต่อการประเมินคุณภาพภายในของคณะกรรมการประเมินคุณภาพภายในต่อไป

3. ควรพิจารณาจัดให้มีการประเมินอภิมานตามความเหมาะสม ของการประเมิน โดยสามารถพิจารณาจัดประเภทการประเมินอภิมานได้ 3 ประเภท ดังต่อไปนี้

1) การประเมินอภิมานด้วยตนเอง (self meta-evaluation) เพื่อเป็นการตรวจสอบคุณภาพของรายงานการประเมิน สำหรับการประเมินคุณภาพภายใน ในระดับสาขาหรือคณะวิชา หรือตรวจสอบรายงานการประเมินอภิมานโดยใช้เทคนิคการประเมินอภิมานสำหรับประเมินตนเอง (meta evaluation for self evaluation) (Yoo, 2007) เพื่อทบทวนรายงานการประเมินที่จัดทำขึ้นด้วยตนเอง หรือตรวจสอบกันเองภายในสาขาวิชา หรือคณะวิชา ให้เกิดความถูกต้อง สมบูรณ์มากขึ้น

2) การประเมินอภิมานภายใน (internal meta-evaluation) เพื่อเป็นการตรวจสอบไขว้คุณภาพของรายงานการประเมินคุณภาพภายใน ระหว่างคณะวิชา หรือภายในสถาบัน โดยใช้นักประเมินอภิมานภายใน ซึ่งสามารถเข้าใจบริบทภายในของสถานศึกษาเป็นอย่างดี สามารถทำการประเมินอภิมานก้าวหน้า (formative meta-evaluation) เพื่อนำผลการประเมินอภิมานมาปรับปรุงแก้ไข การประเมินให้สมบูรณ์มากยิ่งขึ้น เป็นประโยชน์ต่อสถาบันและหน่วยงานต้นสังกัดหรือสำนักงานคณะกรรมการการอุดมศึกษา ทั้งนี้ควรให้นักประเมินประมาณ 2 คนขึ้นไป เพื่อตรวจสอบความสอดคล้องของผลการประเมินอภิมานด้วย

3) การประเมินอภิมานภายนอก (external meta-evaluation) เพื่อเป็นการตรวจสอบคุณภาพของงานประเมินคุณภาพภายนอก โดยสามารถดำเนินการได้ 2 แนวทาง คือ ประเมินอภิมานได้ทั้งนักประเมินภายในและนักประเมินภายนอก ที่ไม่ได้มีส่วนเกี่ยวข้องกับหน่วยงานที่ประเมินนั้น เพื่อเป็นการตรวจสอบกระบวนการประเมิน และสร้างความน่าเชื่อถือของผลการประเมิน นอกจากนี้สามารถทำการประเมินหลังจากเสร็จสิ้นการประเมินแล้ว หรือประเมิน

แบบสรุปรวม (summative meta-evaluation) ทั้งนี้ต้องใช้นักประเมินที่มีประสบการณ์ มีความรู้ความสามารถในการประเมิน เพื่อชี้ให้เห็นจุดแข็ง จุดอ่อนของการประเมินได้เป็นอย่างดี จะเป็นประโยชน์ต่อทั้งหน่วยงานต้นสังกัด สำนักงานคณะกรรมการการอุดมศึกษา และสำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน) ทั้งนี้การกำหนดจำนวนนักประเมิน ควรพิจารณาตามความเหมาะสมและคุณสมบัติ ความน่าเชื่อถือ ความเชี่ยวชาญเฉพาะของนักประเมินเป็นหลัก หากนักประเมินมีความเชี่ยวชาญเฉพาะและมีประสบการณ์ในการประเมินสูง อาจใช้นักประเมิน 1 คน แต่หากนักประเมินมีประสบการณ์การประเมินไม่มากนัก แต่มีความรู้เกี่ยวกับการประเมินควรใช้นักประเมินอย่างน้อย 2 คนขึ้นไป จึงจะเหมาะสมมากกว่า

ข้อเสนอแนะสำหรับการวิจัยในครั้งต่อไป

1. ควรมีการพัฒนา ปรับปรุงค่าน้ำหนักในแบบประเมินอภิมานดังต่อไปนี้ มาตรฐานความตรงของการประเมินร้อยละ 30 ธรรมชาติของร้อยละ 30 จริยธรรมในการประเมิน ร้อยละ 10 ความน่าเชื่อถือของนักประเมินร้อยละ 25 ส่วนประสิทธิผลต้นทุน ร้อยละ 5 เนื่องจากเป็นค่าที่ได้จากการวิเคราะห์หึ่งค์ประกอบเชิงยืนยันจากข้อมูลเชิงประจักษ์ และปรับปรุงคู่มือการประเมินอภิมาน ตัวบ่งชี้ และเกณฑ์ให้เหมาะสม สอดคล้องกับการประเมินรายงานผลการประเมินคุณภาพภายในมากขึ้น
2. ควรทำการประเมินรายงานผลการประเมินคุณภาพภายใน ระดับอุดมศึกษา ในปีการศึกษา 2551 ขึ้น เนื่องจากในการวิจัยใช้รายงานผลการประเมินคุณภาพภายใน ปีการศึกษา 2550 ซึ่งเป็นปีแรกที่กำหนดให้มีการประเมินคุณภาพภายใน ระดับสถาบันที่มีการเผยแพร่ผลการประเมินสู่สาธารณชน และมีกรรมการจากภายนอกสถาบันมาร่วมประเมิน ซึ่งยังพบว่า มีข้อผิดพลาดและมีความไม่สมบูรณ์ในรายงานหลายส่วน
3. ควรนำมาตรฐานการประเมินอภิมาน ไปพัฒนาพัฒนาตัวบ่งชี้ และเกณฑ์ในการประเมินอภิมานงานประเมินอื่นๆ ที่ไม่ใช่รายงานผลการประเมินคุณภาพภายใน เพื่อเป็นการตรวจสอบยืนยันคุณภาพของมาตรฐานการประเมินอภิมานที่พัฒนาขึ้น

รายการอ้างอิง

ภาษาไทย

- เก็จกนก เอื้อวงศ์. 2546. การวิเคราะห์ปัจจัยที่ส่งเสริมและปัจจัยที่เป็นอุปสรรคต่อการดำเนินการในระบบการประกันคุณภาพการศึกษาของสถานศึกษาขั้นพื้นฐาน. วิทยานิพนธ์ปริญญา ดุษฎีบัณฑิต, สาขาวิชาบริหารการศึกษา ภาควิชาบริหารการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- แคลเลีย ศรีแปลก. 2544. ปัจจัยที่ส่งผลต่อการใช้ผลการประเมินภายในสถานศึกษาของโรงเรียน ประถมศึกษา: กรณีศึกษาโรงเรียนคันทนาเยาว กรุงเทพมหานคร. วิทยานิพนธ์ปริญญา มหาบัณฑิต, สาขาวิชาวิจัยการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ทบวงมหาวิทยาลัย. 2539. ข้อมูลและสารสนเทศอุดมศึกษา ทบวงมหาวิทยาลัย. กรุงเทพมหานคร: สำนักนโยบายและแผนอุดมศึกษา สำนักงานปลัดทบวงมหาวิทยาลัย.
- นงลักษณ์ วิรัชชัย. 2550. การประเมินอภิमान: แนวคิดและหลักการ. ใน สุวิมล ว่องวานิช (บรรณาธิการ), การประเมินอภิमान วิธีวิทยาและการประยุกต์ใช้, หน้า 3-20. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- นงลักษณ์ วิรัชชัยและ วรณี เจตจำนงสุข. 2548. การสังเคราะห์รายงานผลการประเมินคุณภาพภายนอกระดับอุดมศึกษา. กรุงเทพมหานคร: สำนักงานรับรองและประเมินคุณภาพการศึกษา (องค์การมหาชน).
- ภัทรารวรรณ ภิรมย์วงศ์. 2546. การวิเคราะห์และประเมินอภิமானการประเมินผลการปฏิบัติงานของนักเรียนในวิชาศิลปศึกษา. วิทยานิพนธ์ปริญญา มหาบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- รติกร ชันดียานันท์. 2548. การศึกษาความสัมพันธ์ระหว่างคุณภาพการประเมินภายใน รูปแบบการใช้ผลการประเมินภายในและผลการประเมินภายนอกของสถานศึกษาขั้นพื้นฐาน. วิทยานิพนธ์ปริญญา มหาบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- รุ่งอรุณ สุเวชวัฒนกุล. 2543. การพัฒนาโมเดลผลการประกันคุณภาพการอุดมศึกษา คณะครุศาสตร์ศึกษาศาสตร์ สังกัดทบวงมหาวิทยาลัย. วิทยานิพนธ์ปริญญา มหาบัณฑิต, สาขาวิชาวิจัยการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

- วราภรณ์ บุญเยี่ยม. 2546. ปัจจัยที่ส่งผลต่อการใช้ผลการประเมินภายนอกเพื่อพัฒนาคุณภาพของสถานศึกษาขั้นพื้นฐาน. วิทยานิพนธ์ปริญญาามหาบัณฑิต, สาขาวิชาวิจัยการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- วสันต์ อดิศักดิ์ และคณะ. 2549. การพัฒนามาตรฐานแห่งชาติทางเทคโนโลยีการศึกษาสำหรับสถาบันผลิตบัณฑิตทางการศึกษา. วารสารสงขลานครินทร์ ฉบับสังคมศาสตร์และมนุษยศาสตร์ 12(4): 455-482.
- วารุณี ลัทธิโชคดี. 2550. การพัฒนาตัวบ่งชี้ร่วมการประเมินคุณภาพสถาบันอุดมศึกษา: การวิเคราะห์ความไม่แน่นอนและความไว. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยและจิตวิทยาการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. 2547ก. ทฤษฎีการประเมิน. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. 2547ข. สถิติประยุกต์สำหรับการวิจัย. พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. 2548. ทฤษฎีการทดสอบแบบดั้งเดิม. พิมพ์ครั้งที่ 5. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ศิริชัย กาญจนวาสี. 2545. ทฤษฎีการทดสอบแนวใหม่. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- บริหารและพัฒนาองค์ความรู้ (องค์การมหาชน), สำนักงาน. ศูนย์ส่งเสริมและพัฒนาพลังแผ่นดินเชิงคุณธรรม (ศูนย์คุณธรรม). 2551. รายงานการประเมินอภิमानภายนอก โครงการเร่งสร้างคุณลักษณะที่ดีของเด็กและเยาวชนไทย. กรุงเทพมหานคร: พริกหวานกราฟฟิค.
- สุภัตรา ทรัพย์เสถียร. 2546. การพัฒนาโมเดลเชิงสาเหตุของประสิทธิผลการประกันคุณภาพภายในสถานศึกษา ระดับการศึกษาขั้นพื้นฐาน: ประยุกต์ใช้โมเดลสมการโครงสร้างกลุ่มพหุ. วิทยานิพนธ์ปริญญาามหาบัณฑิต, สาขาวิชาวิจัยการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- สุวิมล ว่องวานิช. 2550. ประเภท ขั้นตอนของการประเมินอภิमान และคุณสมบัติของนักประเมินอภิमान. ใน สุวิมล ว่องวานิช (บรรณาธิการ), การประเมินอภิमान วิธีวิทยาและการประยุกต์ใช้, หน้า 21-47. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

สมหวัง พิธิยานุวัฒน์. 2549. *วิธีวิทยาการประเมินศาสตร์แห่งคุณค่า*. พิมพ์ครั้งที่ 3.

กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

คณะกรรมการการอุดมศึกษา, สำนักงาน. 2550. (ร่าง) *คู่มือการประกันคุณภาพการศึกษาภายในสถานศึกษาระดับอุดมศึกษา ของคณะกรรมการการอุดมศึกษา*. เอกสารประกอบการประชุมสัมมนา เรื่องการประกันคุณภาพการศึกษา ระดับอุดมศึกษา(8 มีนาคม 2550),. ณ โรงแรมมิราเคิลแกรนด์บอลรูม กรุงเทพมหานคร.

คณะกรรมการการอุดมศึกษา, สำนักงาน. 2552. *คู่มือการประกันคุณภาพการศึกษาภายในสถานศึกษา ระดับอุดมศึกษา*. 1,000 เล่ม. พิมพ์ครั้งที่ 6. กรุงเทพมหานคร: ภาพพิมพ์.

รับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน), สำนักงาน. 2547ก. *หลักเกณฑ์และวิธีการประเมินคุณภาพภายนอกของสถานศึกษาระดับการศึกษาขั้นพื้นฐาน ฉบับปรับปรุง 2547*. กรุงเทพมหานคร: สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา(องค์การมหาชน).

รับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน), สำนักงาน. 2547ข.

พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ. 2547 แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ. 2545.

กรุงเทพมหานคร: สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน).

รับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน), สำนักงาน. 2550ก. *คู่มือการประเมินคุณภาพภายนอกระดับอุดมศึกษา รองที่สอง (พ.ศ. 2549- 2553)*.

กรุงเทพมหานคร: สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน).

รับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน), สำนักงาน. 2550ข. *คำอธิบาย*

ประกอบการใช้แบบประเมินอภิมาน รายงานการประเมินคุณภาพภายนอก

ระดับอุดมศึกษา. 400 เล่ม. กรุงเทพมหานคร: จุดทอง.

รับรองมาตรฐานและประเมินคุณภาพการศึกษา (องค์การมหาชน), สำนักงาน. 2551. *คู่มือการ*

ประเมินผลการปฏิบัติราชการ ตามคำรับรองการปฏิบัติราชการของสถาบันอุดมศึกษา

ประจำปีงบประมาณ พ.ศ. 2551 (ฉบับปรับปรุงครั้งที่ 1)[ออนไลน์]. แหล่งที่มา:

http://www.onesqa.or.th/th/download/check_count.php?FileID=2037&SystemModuleKey=227 [2551, มิถุนายน 25].

- อภิธีร์ ทรงบัณฑิตย์. 2550. *การพัฒนาแบบประเมินสถานศึกษาตามแนวการประกันคุณภาพ การศึกษา ระบบมัลคอม บัลดริจ สำหรับสถานศึกษา ระดับการศึกษาขั้นพื้นฐาน*. วิทยานิพนธ์ปริญญาคุษฎีบัณฑิต, สาขาวิชาการบริหารการศึกษา ภาควิชาบริหารการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ.
- อมรรัตน์ พันธุ์งาม. 2543. *การพัฒนากระบวนการประเมินเพิ่มผลงานและการใช้ประโยชน์จาก ผลการประเมินเพิ่มผลงาน: การประยุกต์ใช้การประเมินอภิมาน*. วิทยานิพนธ์ปริญญา มหาบัณฑิต, สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- อมรวิชัย นาคทรรพ. 2541. *รายงานวิจัย เรื่อง ในกระแสแห่งคุณภาพ*. พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: ที พี พีรินทร์.

ภาษาอังกฤษ

- African Evaluation Association Secretariat. 2002. *The African Evaluation Guidelines: 2002*[Online]. Available From: <http://www.afrea.org/content/index.cfm?navID=5&itemID=204> [2008, June 24].
- American Evaluation Association. 1995. *Guiding Principles for Evaluators*[Online]. Available From: <http://www.preval.org/documentos/00478.pdf> [2007, March 6].
- American Evaluation Association. 2003. *Guiding Principles for Evaluators*[Online]. Available From: <http://www.eval.org/Publications/GuidingPrinciplesPrintable.asp> [2010, March 20].
- Archwamety, T. 2006. *Accreditation and the Academic Program Review*. Kearney: University of Nebraska.
- Berends & Roberts. 2003. Evaluation standards and their application to Indigenous programs in Victoria, Australia. *Evaluation journal of Australasia* 3 new series(2): 54-59.
- Beywl, W. 2003. *Selected Comments to the Standards for Evaluation of the German Evaluation Society*. German Evaluation Society.

- Bjornklide, T. & Bason, C. 2000. *Meta evaluation of the evaluations of higher education in Denmark*[online]. Available From: http://evaluationcanada.ca/distribution/20001012_bjornkilde_thomas_bason_christian.pdf [2008, July 9].
- Bornmann, L., Mittag, S & Daniel, H.-D. 2006. Quality assurance in higher education- meta-evaluation of multi-stage evaluation procedures in Germany. *Higher Education* 52: 687-709.
- Canadian Evaluation Society. 1999. *Essential Skills Series in evaluation*[Online]. Available From: http://evaluationcanada.ca/site.cgi?s=3&ss=31&_lang=en [2007, March 6].
- Coryn, CL. S., Hattie, J.A., Scriven, M. & Hartmann, D.J. 2007. Models and Mechanisms for Evaluating Government-Funded Research An International Comparison. *American Journal of Evaluation* 28(4): 28(4). Available From: <http://aje.sagepub.com> [2008, October 30].
- Coryn, CL. S., & Scriven, M. 2007. Are National-Level Research Evaluation Models Valid, Credible, Useful, Cost-Effective, and Ethical?. *Journal of MultiDisciplinary Evaluation* 4(8): 92-96.
- DanChurchAid. 2007. *Standard format for evaluation report*[online]. Available From: http://www.danchurchaid.org/what_we_do/monitoring_evaluation/Standard_format_for_evaluation_reports.pdf [2010, March 30].
- Davidson, E.J. 2005. *Evaluation methodology basics: the nuts and bolts of sound evaluation*. Thousand Oaks: Sage Publications.
- Ex-detainees Rehabilitation Program. 2002. *Format for Evaluation Report*[online]. Available From: http://www.edrp.gov.ps/Evaluation_Files/Tendering/6-Evaluation%20Report%20Structure.pdf[2010, March 9].
- German Evaluation Society. 2003. *Selected comments to the standards for on the German Evaluation Society: English version*[online]. Available From: http://www.evaluation.org.uk/Pub_library/German%20Evaluation%20Standards_030409.pdf [2008, July 5].

- Ghere, G., King, J.A., Stevahn, L. S. & Minnema, J. 2006. A professional Development Unit for Reflecting on Program Evaluator Competencies. *American Journal of Evaluation* 27(1): 108-123.
- Griffin, N.C., and Miyoshi, J.N. 2009. *Third Year Report: Evaluation of the Artful learning Program*. Los Angeles: University of California.
- Jang, S. 2000. The Appropriateness of Joint Committee Standards in Non-Western Settings: A Case Study of South Korea. In Craig Russon ed. (The Evaluation Center), *The Program Evaluation Standards in International Settings*, 41-59.
- Joint Committee on Standard for Educational Evaluation. 1994. *The Program Evaluation Standards*[online]. Available From: <http://www.wmich.edu/evalctr/jc/PGMSTNDS-SUM.htm> [2008, May 26].
- Joint Committee on Standard for Educational Evaluation. 2010. *The Program Evaluation Standards*[online]. Available From: <http://www.jcsee.org/program-evaluation-standards> [2010, April 7].
- Laubli Loud, M.M. 2004. *Setting standards and providing guideline* [online]. Available From: <http://www.bag.admin.ch/evaluation/03124/> [2008, July 9].
- Lei, P., Smith, M., and Suen, H.K. 2007. The Use of Generalizability Theory to Estimate Data Reliability in Single-Subject Observational Research. *Psychology in the Schools* Vol. 44(5): 433-439.
- Lin, C., Lin, B.B., and Chern, H. 2009. APEC Pharmaceutical Evaluation Report (PER) Scheme: A Regional Solution for Global Drug Development. *Drug Information Journal* May 43 (3): 305-309.
- Madaus, G. & Mcsherry, J. 2007. *Second Update on a Methodology for Assessing the Effectiveness of RCAB Programs for Child Protection and Abuse Prevention*[online]. Available From: <http://www.rcab.org/childAdvocacy/EffectivenessAssessment-070331.pdf> [2009, January 31].

- Migotsky, C. & Stake, R. 2001. *An Evaluation of an Evaluation: CIRCE's Metaevaluation of the Site Visits and Issue Papers of the ATE Program Evaluation*[online]. Available From: <http://www.wmich.edu/evalctr/ate/sitevisitmetafinal.pdf> [2008, February 2].
- Oxfam Australia. 2006. *Meta-Evaluation of Evaluation Studies Completed by Oxfam Australia between January 2002-June 2006*[online]. Available From: <http://www.oxfam.org.auidocuments/PDF/meta-evaluation.pdf> [2009, January 31].
- Patel, M. 2002. A meta-evaluation, or quality assessment, of the evaluations in this issue, based on African Evaluation Guidelines: 2002. *Evaluation and Program Planning* 25: 329-332.
- Patel, M. 2002. *The African Evaluation Guidelines: 2002*[online]. Available From: <http://www.afrea.org/content/index.cfm?navID=5&itemID=204> [2008, June 24].
- Russon, C & Koehly, L.M. 1995. Construction of a Scale to Measure the Persuasive Impact of Qualitative and Quantitative Evaluation Reports. *Evaluation and Program Planning*. 18(2): 165-177.
- Scott-little, C., Hamann, M., & Jurs, S.G. 2002. Evaluations of After-School Programs: A Meta-Evaluation of Methodologies and Narrative Synthesis of Findings. *American Journal of Evaluation* 23(4): 387-419.
- Sergers, M. & Dochy, F. 1996. Quality assurance in higher education: Theoretical considerations and empirical evidences. *Studies in Educational Evaluation* 22(2): 115-137.
- Shavelson, R.J., and Webb, N.M. 1991. *Generalizability Theory A Primer*. Newbury park: CA: Sage.
- Sriven, M. 1969. An introduction to meta-evaluation. *Educational Product Report* 2: 36-38.
- Sriven, M. 1991. *Evaluation thesaurus*. Newbury park: CA: Sage.
- Sriven, M. 2003. *Meta-evaluation*[online]. Available from: <http://eval.cgu.edu/lectures/evtypes/metaeval/> [2008, June 9].
- Stufflebeam, D.L. 1974. *Meta-evaluation*[Online]. Available From: <http://www.wmich.edu/evalctr/pubs/ops/ops03.pdf> [2008, February 2].

- Stufflebeam, D.L. 1999. *Program evaluation models meta evaluation checklist*[Online]. Available From: [http://www.wmich.edu/evalctr/checklist/program evaluation checklist.pdf](http://www.wmich.edu/evalctr/checklist/program%20evaluation%20checklist.pdf) [2008, January 27].
- Stufflebeam, D.L. 2000. The methodology of metaevaluation as reflected in metaevaluations by the Western Michigan University Evaluation center. *Journal of Personal Evaluation in Education* 14(1): 95-125.
- Stufflebeam, D.L. 2001. The metaevaluation imperative. *American Journal of Evaluation* 22(2): 183-209.
- Stufflebeam, D.L., Madaus, G.F., & Kellagham, T. 2000. *Evaluation models*. Kluwer Academic Publishers.
- Turner, R.C. & Carlson, L. 2003. Indexes of Item-Objective Congruence for Multidimensional Items. *International Journal of Testing* 3(2): 163-171.
- United Nations Development Programme (UNDP). 2008. *Evaluation Report*[Online]. Available From: <http://www.undp.org/evaluation/methodologies.htm>[2010, January 3].
- United States Agency International Development (USAID). 2006. *Constructing An Evaluation Report*[Online]. Available From: <http://www.usaid.gov/policy/evalweb/>[2010, January 3].
- Victoria University of Wellington. 2005. *Report format*[Online]. Available From: http://www.victoria.ac.nz/st_services/slss/studyhub/assignment/Report-format.pdf[2010, March 9].
- Widmer, T., Landert, C. & Bachmann, N. 2000. *Evaluation Standards of SEVAL, the Swiss Evaluation Society (SEVAL Standards)*[online]. Available From: http://www.evaluation.org.uk/Pub_library/SwissEvaluationStandards_2000_E.pdf [2008, July 5].
- Windsor, R.A., Boyd, N.R., & Orleans, C.T. 1998. A meta-evaluation of smoking cessation intervention research among pregnant women: improving the science and art. *Health Education Research Theory & Practice* 13(3): 419-438.

Yoo, S.J. (R&D Evaluation Center). 2007. *National Evaluation System of Public R&D*

Program in Korea (kNES)[PowerPoint]. Seminar on Evaluating Innovative

Approaches to Public Service Delivery Madrid (Spain): Korea Institute of S&T

Evaluation and Planning.



ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ก
รายนามผู้เชี่ยวชาญ
รายนามผู้ทรงคุณวุฒิ

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ก
รายนามผู้เชี่ยวชาญด้านการประกันคุณภาพ และการวัดผล

ที่	ชื่อ	ตำแหน่ง
1	ผู้ช่วยศาสตราจารย์ ดร. จุมพล พูลภัทรชีวิน	ผู้ช่วยศาสตราจารย์ ภาควิชานโยบาย การจัดการ และความเป็นผู้นำทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2	ดร. คมศร วงษ์รักษา	รองผู้อำนวยการ สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา
3	นาวาตรีหญิง ดร. กิตติยา เอ็ฟฟานส	รองผู้อำนวยการ สำนักงานรับรองมาตรฐานและประเมินคุณภาพการศึกษา
4	รองศาสตราจารย์ ดร. ณรงค์ อยู่ถนอม	รองอธิการบดี มหาวิทยาลัยศรีปทุม
5	ดร. วราภรณ์ ไทยมา	ผู้อำนวยการหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาบริหารการศึกษา มหาวิทยาลัยศรีปทุม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายนามผู้ทรงคุณวุฒิด้านการประกันคุณภาพ ระดับอุดมศึกษา

ที่	ชื่อ	สังกัด/หน่วยงาน
1	ศาสตราจารย์กิตติคุณ ดร.นงลักษณ์ วิรัชชัย	อาจารย์พิเศษภาควิชาวิจัยและจิตวิทยา การศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
2	รองศาสตราจารย์ เทือน ทองแก้ว	ประธานคณะกรรมการประกันคุณภาพ ประจำมหาวิทยาลัยราชภัฏสวนดุสิต
3	ผู้ช่วยศาสตราจารย์ ดร.พันธ์ศักดิ์ พลสารมัย	รองคณบดีฝ่ายแผน งบประมาณ และประกันคุณภาพ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
4	ดร.ไสมฉาย บุญญานันต์	อาจารย์ประจำภาควิชาศิลปะ ดนตรี และ นาฏศิลป์ศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
5	ดร. วราภรณ์ ไทโยมา	ผู้อำนวยการหลักสูตรปรัชญาดุษฎีบัณฑิต สาขาบริหารการศึกษา มหาวิทยาลัยศรีปทุม
6	ผู้ช่วยศาสตราจารย์ ดร.กาญจนา มีศิลป์วิภักย์	คณบดีคณะนิเทศศาสตร์ มหาวิทยาลัยศรีปทุม
7	ผู้ช่วยศาสตราจารย์เจิดฉิน สุขปลั่ง	คณบดีคณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยราชภัฏสวนดุสิต
8	ผู้ช่วยศาสตราจารย์ ดร.วิรัช เลิศไพฑูรย์พันธุ์	ผู้ช่วยอธิการบดีด้านวิชาการ มหาวิทยาลัยศรีปทุม
9	รองศาสตราจารย์ จันทนา ทองประยูร	อาจารย์ประจำสาขานิเทศศาสตร์ มหาวิทยาลัยสุโขทัยธรรมมาธิราช
10	ดร. คมศร วงษ์รักษา	รองผู้อำนวยการ สำนักงานรับรองมาตรฐาน และประเมินคุณภาพการศึกษา



ภาคผนวก ข
เครื่องมือในการวิจัย

ศูนย์วิจัยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ข เครื่องมือในการวิจัย

คู่มือประเมินอภิमान

1. แนวคิดและความเป็นมาของหลักสูตรฝึกอบรมนักประเมินอภิमान

สำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) ได้พัฒนาหลักสูตรอบรมผู้ประเมินคุณภาพการศึกษาภายใน เพื่อพัฒนาบุคลากรของระบบอุดมศึกษาให้มีความรู้ ความเข้าใจ เกี่ยวกับการดำเนินงานประกันคุณภาพการศึกษาที่ได้พัฒนาขึ้น รวมทั้งสร้างผู้ประเมินคุณภาพการศึกษาภายในให้มีแนวปฏิบัติในการประเมินคุณภาพการศึกษาภายในที่เป็นมาตรฐานเดียวกัน และทำหน้าที่ส่งเสริม พัฒนาการดำเนินงานประกันคุณภาพการศึกษาภายในสถาบันอุดมศึกษา ให้มีความต่อเนื่องและยั่งยืน มีประสิทธิภาพและประสิทธิผลยิ่งขึ้น อีกทั้งสร้างความมั่นใจในการเตรียมความพร้อมรับการประเมินคุณภาพภายนอก และสร้างวัฒนธรรมคุณภาพให้เกิดขึ้นในสถาบันอุดมศึกษา

นอกจากนี้ สกอ. ได้จัดให้สถาบันอุดมศึกษา ได้รับการประเมินคุณภาพภายใน ระดับสถาบัน จากผู้ประเมินคุณภาพภายใน ของสกอ. ที่อย่างน้อยร้อยละ 50 มาจากภายนอกสถาบัน เพื่อเป็นการควบคุมคุณภาพการประเมินคุณภาพภายใน ทั้งนี้เพื่อเป็นการตรวจสอบซ้ำจากรายงานผลการประเมินคุณภาพภายใน ระดับสถาบัน ให้มีคุณภาพน่าเชื่อถือ จึงมีความจำเป็นในการควบคุมคุณภาพของการประเมินคุณภาพภายในให้เป็นมาตรฐานเดียวกัน ด้วยการประเมินอภิमान (Meta-evaluation) เพื่อตัดสินคุณค่าของรายงานผลประเมินเทียบกับเกณฑ์มาตรฐานที่พัฒนาขึ้น เพื่อวิเคราะห์หาจุดแข็ง จุดอ่อนในการดำเนินงาน และกำกับติดตามการนำผลการประเมินไปใช้ประโยชน์ของสถาบัน จึงได้พัฒนาหลักสูตรอบรมนักประเมินอภิमानภายนอก โดยใช้ผู้ประเมินคุณภาพภายในที่ผ่านการอบรมผู้ประเมินคุณภาพภายใน จากสกอ. มีความรู้ ความเข้าใจในมาตรฐาน ตัวบ่งชี้ และกระบวนการประเมินคุณภาพภายใน เพื่อทำหน้าที่เป็นนักประเมินอภิमान ในระดับอุดมศึกษา เพื่อทำหน้าที่ควบคุม ตรวจสอบ และรับรองผลการประเมินคุณภาพภายใน ที่ผ่านการประเมินจากคณะกรรมการประเมินคุณภาพภายใน ของสถาบัน

2. วัตถุประสงค์ของหลักสูตร

1. เพื่อสร้างความรู้ ความเข้าใจ และบทบาทหน้าที่ของนักประเมินอภิमान
2. เพื่อให้ นักประเมินอภิमानมีทักษะและประสบการณ์ตรงในการประเมินอภิमानจริง
3. เพื่อให้ นักประเมินอภิमानสามารถนำความรู้ที่ได้รับไปพัฒนาส่งเสริมให้มีการประเมินอภิमानภายในหน่วยงานของตนเองและหน่วยงานที่เกี่ยวข้อง

3. มาตรฐานการประเมินอภิมาน

การประเมินอภิมานประกอบด้วยมิติการประเมิน 5 มาตรฐาน 38 ตัวบ่งชี้ โดยมีคำอธิบาย รายละเอียดและข้อมูลดังต่อไปนี้

มาตรฐานที่ 1 ความตรงของการประเมิน (Validity)

คุณลักษณะของงานประเมิน ที่มีการจัดการเอกสารในการประเมิน การวิเคราะห์บริบท กำหนดวัตถุประสงค์และออกแบบการประเมิน การเก็บรวบรวมสารสนเทศ วิเคราะห์ แปลผลและสรุปสารสนเทศอย่างมีคุณภาพ และตรวจสอบการประเมินอย่างถูกต้องสมเหตุสมผล ประกอบด้วย 13 ตัวบ่งชี้ คือ

Va.1.1 การระบุบริบทของสถาบันที่ส่งผลกระทบต่อประเมิน

คำอธิบาย การแสดงถึงประวัติ ภูมิหลัง ข้อมูลพื้นฐานของสถาบัน การดำเนินงาน และผลการดำเนินงาน รวมถึงสภาพแวดล้อมของสถาบันที่มีอิทธิพลและการวิเคราะห์ผลกระทบต่อประเมิน ได้แก่ นโยบายการจัดการศึกษา แผนยุทธศาสตร์อุดมศึกษาระยะยาว 15 ปี สภาพสถาบัน สภาพแวดล้อมทางเศรษฐกิจ สังคม การเมือง คุณลักษณะข้อมูลที่เกี่ยวข้องกับระบบประเมินคุณภาพ สภาพการแข่งขันของรัฐและเอกชน เป็นต้น

วิธีการตรวจสอบ

- มีการระบุข้อมูลพื้นฐาน ของสถาบันที่ประเมิน
- มีการระบุข้อมูลประวัติความเป็นมาของสถาบันที่ประเมิน
- มีการบรรยายถึงบริบทที่แวดล้อมสถาบันที่ประเมิน

เกณฑ์การประเมิน

1. ไม่ระบุประวัติหรือข้อมูลพื้นฐาน ของสถาบันในรายงานการประเมิน
2. มีการระบุประวัติและข้อมูลพื้นฐาน ของสถาบันในรายงานการประเมิน
3. มีการระบุประวัติ ข้อมูลพื้นฐาน และมีการอธิบายถึงสภาพทั่วไป การดำเนินงาน และผลการดำเนินงาน ของสถาบันที่ประเมินอย่างชัดเจน
4. มีการระบุประวัติ ข้อมูลพื้นฐาน สภาพทั่วไป การดำเนินงาน และผลการดำเนินงาน ของสถาบันที่ประเมินอย่างชัดเจน รวมถึงวิเคราะห์ผลกระทบต่อดำเนินการ และผลผลิตของสถาบัน

Va.1.2 การระบุถึงอัตลักษณ์ของสถาบัน

คำอธิบาย อัตลักษณ์ของสถาบัน เป็นลักษณะเด่นของสถาบัน ที่เป็นที่ยอมรับของสาธารณชน โดยสถาบันมีการกำหนดนโยบาย และแผนการดำเนินงาน เพื่อให้สามารถบรรลุจุดเน้นหรืออัตลักษณ์ของสถาบันได้

วิธีการตรวจสอบ

- มีการระบุอัตลักษณ์หรือจุดเน้น ของสถาบันที่ประเมิน
- มีการวางแผนการประเมินตามอัตลักษณ์หรือจุดเน้นสถาบัน
- มีผลการประเมินตามอัตลักษณ์หรือจุดเน้นของสถาบัน

เกณฑ์การประเมิน

1. ไม่มีการระบุถึงอัตลักษณ์หรือจุดเน้นของสถาบันไว้ในรายงานการประเมิน
2. มีการระบุถึงอัตลักษณ์หรือจุดเน้นของสถาบันไว้ในรายงานการประเมิน
3. มีการระบุถึงอัตลักษณ์หรือจุดเน้นของสถาบันและมีการกำหนดนโยบาย แผนการประเมินตามจุดเน้นดังกล่าว
4. มีการระบุถึงอัตลักษณ์หรือจุดเน้นของสถาบันและมีกิจกรรมดำเนินการตามแผน เพื่อให้สามารถบรรลุจุดเน้นดังกล่าว

Va1.3 การบรรยายถึงวัตถุประสงค์ในการประเมินอย่างชัดเจน

คำอธิบาย วัตถุประสงค์ในการประเมิน เป็นการแสดงจุดหมายของการประเมินอย่างชัดเจน เพื่อให้ผู้มีส่วนได้ส่วนเสียสามารถตรวจสอบได้ว่า การประเมินได้ครอบคลุมนโยบาย การพัฒนาการประเมิน รวมถึงข้อตกลงเบื้องต้นของปัจจัยภายนอกที่มีส่วนในการวางแผนการประเมินในสามารถบรรลุวัตถุประสงค์ได้

วิธีการตรวจสอบ

- มีการระบุถึงวัตถุประสงค์หรือเป้าหมายการประเมิน
- มีการวางแผนการประเมินให้บรรลุวัตถุประสงค์
- มีการระบุแนวทางการประเมิน เพื่อให้บรรลุวัตถุประสงค์

เกณฑ์การประเมิน

1. ไม่มีการระบุวัตถุประสงค์ ในการประเมินไว้ในรายงาน
2. มีการระบุวัตถุประสงค์ทั่วไปในการประเมินไว้ในรายงานการประเมิน
3. มีการระบุวัตถุประสงค์ในการประเมินอย่างชัดเจน ไว้ในรายงานการประเมิน
4. มีการระบุถึงวัตถุประสงค์เฉพาะไว้ในรายงาน อย่างชัดเจน สามารถตรวจสอบได้ และอธิบายแนวทางการประเมินเพื่อให้บรรลุวัตถุประสงค์

Va1.4 การออกแบบการประเมินที่สอดคล้องกับสภาพ และบริบทของสถาบัน

คำอธิบาย การออกแบบการประเมิน เป็นการบรรยายลักษณะของการประเมิน ที่น่าเชื่อถือ การแสดงหลักฐานสนับสนุนคุณภาพการประเมิน ตั้งแต่วิธีการเก็บรวบรวมข้อมูล ได้แก่ การเลือกผู้ตัวอย่าง เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ชนิดของข้อมูล ตลอดจนเทคนิคการวิเคราะห์ข้อมูล ช่วงเวลาและทรัพยากรที่จำเป็น โดยมีการพิจารณาถึงผลกระทบต่อข้อค้นพบและ

ข้อสรุปในการประเมิน และข้อจำกัดในการประเมิน นอกจากนี้มีการออกแบบการเขียนรายงานการประเมินที่สอดคล้องกับสภาพและบริบทของสถาบัน

วิธีการตรวจสอบ

- มีการกำหนดแผนการประเมินไว้ล่วงหน้า โดยมีร่องรอยว่ามีการประชุมเพื่อวางแผนการประเมิน ก่อนการประเมิน
- มีการกำหนดภาระหน้าที่ ในการประเมิน การรวบรวมข้อมูล วิเคราะห์ข้อมูล การแปลความหมาย และการเขียนรายงาน ในทีมผู้ประเมิน
- แผนการประเมินมีความยืดหยุ่น ในการบริหารจัดการ กรณีที่มีมุมมองแตกต่างกัน

เกณฑ์การประเมิน

1. ไม่มีรายละเอียดขั้นตอนเตรียมการวางแผนการประเมิน
2. มีการวางแผนและออกแบบการประเมินเฉพาะ การเก็บรวบรวมข้อมูล หรือการวิเคราะห์ข้อมูลไว้ ก่อนการประเมิน
3. มีการวางแผน ออกแบบ การประเมิน ได้แก่ แหล่งข้อมูล การเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล การแปลความหมายและเขียนรายงานไว้ในขั้นตอนการเตรียมการและวางแผนก่อนการประเมิน
4. มีการวางแผนและออกแบบการประเมิน ได้แก่ แหล่งข้อมูล การเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล การแปลความหมายและเขียนรายงาน สอดคล้องกับสภาพ บริบทของสถาบัน

Va1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสารในการประเมิน

คำอธิบาย เอกสารในการประเมิน ได้แก่ รายงานการประเมินตนเอง (SAR) เอกสารอ้างอิง ได้แก่ แผนงาน โครงการ/กิจกรรม รายงานผลการดำเนินงาน และหลักสูตร การจัดการเรียนการสอน การประเมินผล รวมถึงผลงานต่างๆ เป็นต้น มีการแสดงรายละเอียดในการพิจารณา วิเคราะห์ ความแตกต่างของเอกสาร หลักฐานจากแหล่งต่างกัน

วิธีการตรวจสอบ

- มีการระบุข้อมูลที่มาหลายแหล่ง ได้แก่ จากการสัมภาษณ์ ศึกษาเอกสารหลักฐาน สังเกตหรือตรวจเยี่ยมพื้นที่ รวมถึงฐานข้อมูล
- มีการอธิบายถึงการวิเคราะห์ความแตกต่างของเอกสาร หรือข้อมูลที่รวบรวมมาได้

เกณฑ์การประเมิน

1. เอกสารในการประเมิน ใช้เพียงแหล่งเดียว คือ รายงานการประเมินตนเอง (SAR)
2. มีรายละเอียดการรวบรวมเอกสารประกอบในการประเมิน และรายงานผลการประเมิน
3. มีรายละเอียดการรวบรวมเอกสารการประเมินชัดเจน ระบุที่มาของเอกสารจากหลายแหล่ง แต่

ไม่มีผลการวิเคราะห์ความแตกต่างของเอกสาร

4. มีรายละเอียดการรวบรวมเอกสารการประเมินอย่างถูกต้องและชัดเจน โดยแสดงผลการวิเคราะห์ความแตกต่างของเอกสารที่มาจากแหล่งต่างกัน และรายงานผลการประเมินอย่างถูกต้องเหมาะสม

Va1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ

คำอธิบาย ข้อมูลที่ใช้ในการประเมินคุณภาพภายใน ต้องมีความน่าเชื่อถือ ได้มาจากการแหล่งการเก็บรวบรวมข้อมูลจริงที่ได้มาตรฐาน จากหลายแหล่ง ทั้งการสังเกตสถานที่ การดำเนินงาน การสัมภาษณ์ ผู้ที่เกี่ยวข้องทุกฝ่าย และการศึกษาจากเอกสารหลักฐานที่เกี่ยวข้อง และการรวบรวมข้อมูลอื่นๆ โดยนำมาเปรียบเทียบเพื่อพิจารณาความสอดคล้อง น่าเชื่อถือของข้อมูล

วิธีการตรวจสอบ

- มีการระบุแหล่งที่เก็บรวบรวมข้อมูล เช่น ฐานข้อมูล หลักฐานเอกสาร การสัมภาษณ์ จากบุคคลหลายกลุ่ม ได้แก่ ผู้บริหาร อาจารย์ นักศึกษา ศิษย์เก่า ผู้ปกครอง ผู้ใช้บัณฑิต เป็นต้น การสังเกตข้อมูล และการตรวจเยี่ยมสถานที่
- มีเครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลที่น่าเชื่อถือ
- มีการระบุวิธีการตรวจสอบข้อมูล เช่น การตรวจสอบแบบสามเส้า จากข้อมูลมียังไม่ชัดเจน

เกณฑ์การประเมิน

1. ใช้แหล่งข้อมูลจากการศึกษาเอกสารเพียงแหล่งเดียวในการเก็บรวบรวมข้อมูล
2. ใช้แหล่งข้อมูลเพียง 2 แหล่ง ได้แก่ การศึกษาเอกสาร การสัมภาษณ์ ในการเก็บรวบรวมข้อมูล และมีการรายงานไว้
3. มีกระบวนการเก็บรวบรวมข้อมูลหลายแหล่งทั้งการตรวจเยี่ยม สัมภาษณ์ และศึกษาเอกสาร
4. มีกระบวนการเก็บรวบรวมข้อมูลหลายแหล่ง โดย ตรวจสอบแบบสามเส้าเปรียบเทียบแหล่งข้อมูลแต่ละแหล่งที่เก็บรวบรวมมาสอดคล้องกับเกณฑ์ มาตรฐานการประเมิน

Va1.7 การตรวจสอบความถูกต้องของข้อมูล

คำอธิบาย ความถูกต้องของข้อมูลในการประเมินคุณภาพภายใน นอกจากข้อมูลการประเมินมาจากหลายแหล่งแล้ว ยังต้องมีการตรวจสอบว่าข้อมูลเหล่านั้นเป็นจริง โดยทำการตรวจสอบซ้ำจากการบันทึกข้อมูล หรือเปรียบเทียบข้อมูลก่อนและหลังการประเมิน เพื่อแสดงหลักฐานที่ยืนยันได้ว่า ข้อมูลมีความสอดคล้องกัน และมีการควบคุมความคลาดเคลื่อนจากปัจจัยที่ส่งผลกระทบต่อคุณภาพของข้อมูล เพื่อความถูกต้องในการตัดสินผลการประเมินตามมาตรฐาน และเกณฑ์การประเมิน

วิธีการตรวจสอบ

- มีการระบุถึงการตรวจสอบความถูกต้องของข้อมูล ด้วยเทคนิคต่างๆ เช่น จากบุคคลหลายกลุ่ม หรือจากหลักฐานเอกสารหลายแหล่ง เป็นต้น
- มีการอธิบายถึงหลักฐานที่ใช้ในการตรวจสอบความถูกต้องและความสอดคล้องของข้อมูล

เกณฑ์การประเมิน

1. ไม่ระบุการตรวจสอบความถูกต้องของข้อมูล
2. มีการระบุการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา แต่ไม่มีหลักฐานยืนยันแสดงความสอดคล้องกันของข้อมูล
3. มีการระบุการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา แต่หลักฐานยืนยันความสอดคล้องกันของข้อมูลต่างๆ ยังไม่มีรายละเอียด
4. มีการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา และมีหลักฐานยืนยันความสอดคล้องกันของข้อมูลต่างๆ มีรายละเอียดชัดเจน

Va1.8 การตรวจสอบคุณภาพของสารสนเทศ ตรงกับความต้องการของผู้ใช้ข้อมูล

คำอธิบาย สารสนเทศในการประเมิน มาจากข้อมูลที่รวบรวมมา มีความพอเพียง และผ่านการวิเคราะห์แล้ว ซึ่งมีความถูกต้อง สมบูรณ์ และน่าเชื่อถือ นอกจากนี้ยังสอดคล้องกับความต้องการของผู้ใช้ข้อมูล

วิธีการตรวจสอบ

- มีการระบุถึงคุณภาพของสารสนเทศที่ได้อบรมรวมมา
- มีการบ่งบอกว่าสารสนเทศที่รวบรวมมาได้ตรงกับความต้องการของผู้ใช้ข้อมูล

เกณฑ์การประเมิน

1. ไม่มีการตรวจสอบคุณภาพของสารสนเทศที่เก็บรวบรวมมาได้
2. มีการตรวจสอบคุณภาพการเก็บรวบรวมข้อมูล สารสนเทศบางส่วน
3. มีการระบุการตรวจสอบสารสนเทศที่เก็บรวบรวมมากับความต้องการของผู้ใช้ข้อมูล
4. มีการตรวจสอบคุณภาพความถูกต้อง สมบูรณ์ของสารสนเทศ สอดคล้องและตรงกับความต้องการของผู้ใช้ข้อมูลส่วนมาก

Va1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูลอย่างถูกต้องตามวิธีการที่กำหนด

คำอธิบาย ตัวบ่งชี้ในการประเมินคุณภาพภายใน มีทั้งตัวบ่งชี้เชิงปริมาณ และเชิงคุณภาพ ซึ่งต้องมีการวิเคราะห์ข้อมูลที่ถูกต้อง ตรวจสอบความถูกต้องได้จากการรวบรวมข้อมูลดิบใน common dataset ของสถาบัน และสูตรในการคำนวณในคู่มือการประกันคุณภาพการศึกษา

ภายในสถานศึกษา ระดับอุดมศึกษาของสกอ. และเปรียบเทียบกับมาตรฐาน เกณฑ์การประเมิน คุณภาพภายใน

วิธีการตรวจสอบ

- มีการระบุวิธีการวิเคราะห์ข้อมูลตามวิธีที่กำหนด ทั้งเชิงปริมาณและเชิงคุณภาพ
- มีการอธิบายถึงหลักฐานยืนยันความสอดคล้องของข้อมูลรายตัวบ่งชี้

เกณฑ์การประเมิน

1. มีการวิเคราะห์ข้อมูลเฉพาะข้อมูลเชิงปริมาณเท่านั้น แต่ยังมีความถูกต้อง **น้อยกว่า** 36 ตัวบ่งชี้
2. การวิเคราะห์ข้อมูลเชิงปริมาณ และเชิงคุณภาพ มีความถูกต้อง 36 ตัวบ่งชี้ ขึ้นไป
3. การวิเคราะห์ข้อมูลเชิงปริมาณ และเชิงคุณภาพ มีความถูกต้องอย่างน้อย 39 ตัวบ่งชี้ และมีหลักฐานยืนยันความสอดคล้องของข้อมูลต่างๆ
4. มีการวิเคราะห์ข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพอย่างเป็นระบบ มีความถูกต้อง น่าเชื่อถือทุกตัวบ่งชี้ สามารถประมวลผลข้อมูลได้อย่างมีประสิทธิภาพ

Va1.10 การแปลผลและสรุปผลการประเมินถูกต้องตามหลักฐานที่ปรากฏ

คำอธิบาย การแปลผลและสรุปผลการประเมิน โดยการตัดสินผลการประเมิน ดีความจาก ข้อค้นพบ เป็นการตัดสินคุณค่าของข้อค้นพบเทียบกับมาตรฐาน เกณฑ์การประเมิน จากข้อมูลเชิง ประจักษ์ (empirical finding) ตามตัวบ่งชี้ และสรุปผลการตัดสินผลการประเมินจากการพิจารณา ร่วมกันของทีมผู้ประเมิน

วิธีการตรวจสอบ

- มีการลงสรุปสอดคล้องกับผลการประเมินและผลการวิเคราะห์ข้อมูล
- มีการอ้างอิงสารสนเทศที่สนับสนุนการประเมิน

เกณฑ์การประเมิน

1. การแปลผลและสรุปผล ไม่สอดคล้องกับผลการประเมิน และยังขาดการอ้างอิงสารสนเทศ ที่สนับสนุนผลการประเมิน
2. การแปลผลและสรุปผลการประเมิน สอดคล้องกับผลการวิเคราะห์ข้อมูล แต่ยังไม่มีการ อ้างอิงสารสนเทศที่สนับสนุนผลการประเมิน
3. การแปลผลและสรุปผลการประเมินถูกต้อง สอดคล้องกับผลการวิเคราะห์ข้อมูล และมีการ อ้างอิงสารสนเทศที่สนับสนุนผลการประเมินบางส่วน
4. การแปลผลและสรุปผลการประเมินอย่างถูกต้อง ชัดเจน มีการอ้างอิงสารสนเทศสนับสนุนผล การประเมินทุกครั้ง และสอดคล้องกับผลการวิเคราะห์ข้อมูล

Va1.11 การเขียนรายงานผล ตามข้อค้นพบทั้งเชิงบวกและเชิงลบ

คำอธิบาย การเขียนรายงานผลการประเมิน ตามข้อค้นพบจากการประเมินและผลการวิเคราะห์ข้อมูล และการสรุปผลการประเมิน ซึ่งความน่าเชื่อถือของข้อสรุปมาจากหลักฐานอ้างอิง และความเชี่ยวชาญของทีมผู้ประเมิน ที่แสดงข้อค้นพบทั้งเชิงบวก และเชิงลบที่ได้จากการประเมิน มีการนำเสนอทั้งจุดเด่น และจุดที่ควรพัฒนาในการประเมินอย่างชัดเจนและสมบูรณ์

วิธีการตรวจสอบ

- มีการระบุถึงจุดเด่น จุดด้อยหรือข้อจำกัดของการประเมิน
- การเขียนรายงานมีความถูกต้อง เสนอข้อค้นพบทั้งเชิงบวกและเชิงลบ

เกณฑ์การประเมิน

1. การเขียนรายงานผลการประเมิน **ไม่มีการ**นำเสนอจุดเด่นและจุดที่ควรพัฒนา ของการประเมิน
2. การเขียนรายงานผลการประเมิน เสนอเฉพาะข้อค้นพบด้านเดียว เป็นเชิงบวก หรือเชิงลบ เท่านั้น
3. การเขียนรายงานผลการประเมินเน้นจุดเด่น แต่มีการนำเสนอทั้งจุดเด่นและแนวทางการพัฒนาการประเมินบางส่วน
4. การเขียนรายงานผลการประเมินอย่างถูกต้อง ทั้งเชิงบวกและเชิงลบ โดยเสนอทั้งจุดเด่นและแนวทางการพัฒนาการประเมินชัดเจน

Va1.12 การรายงานผลการประเมินอย่างเป็นกลาง

คำอธิบาย การรายงานผลการประเมิน มาจากข้อค้นพบ ข้อสรุปจากการประเมินที่ปราศจากการบิดเบือนผลการประเมิน โดยตัวบ่งชี้ในการประเมินมีข้อมูลสนับสนุนเหตุผลในการวิเคราะห์ผลการประเมิน และแสดงให้เห็นว่า มีการพิจารณาตัดสินผลการประเมินจากการมีส่วนร่วมของผู้มีส่วนได้ส่วนเสียทุกกลุ่ม

วิธีการตรวจสอบ

- มีการระบุถึง รายงานผลการประเมินที่เป็นที่ยอมรับของทุกกลุ่ม
- มีการอธิบายถึงการพิจารณา การป้องกัน บิดเบือนรายงานผลการประเมินจากผู้ที่มีอิทธิพลบางกลุ่ม
- รายงานผลการประเมินมีความเป็นกลาง กับผู้ที่เกี่ยวข้องทุกฝ่าย

เกณฑ์การประเมิน

1. **ไม่มี** หลักฐานแสดงว่าผลการประเมินรายงานอย่างเป็นกลาง
2. มีการระบุว่า รายงานผลการประเมินอย่างเป็นกลาง แต่ไม่มีหลักฐานว่าพิจารณาจากการตัดสินใจของกลุ่มผู้ที่เกี่ยวข้องกับการประเมิน

3. มีการรายงานผลการประเมินอย่างเป็นกลาง **ไม่มี**หลักฐานแสดงว่าผลการประเมินถูกบิดเบือน และระบุว่าผู้ที่เกี่ยวข้องกับการประเมิน ส่วนใหญ่ร่วมพิจารณาตัดสิน

4. มีการรายงานผลการประเมินอย่างชัดเจน ตรวจสอบได้ โดยผลการประเมินสามารถรองรับการตัดสินใจของทุกกลุ่มที่เกี่ยวข้อง

Va1.13 การตรวจสอบความถูกต้องของรายงานการประเมินจากผู้ที่เกี่ยวข้อง

คำอธิบาย ผู้ที่เกี่ยวข้องหรือผู้มีส่วนได้ส่วนเสียหลักสามารถเข้ามามีส่วนร่วมในการทบทวนร่างรายงานการประเมินเพื่อตรวจสอบความสอดคล้องของข้อเท็จจริง ข้อผิดพลาด หรือข้อยกเว้น ที่มีผลต่อการตีความข้อค้นพบจากการประเมินที่อาจทำให้เกิดการพิจารณาที่ไม่ถูกต้อง ผู้ประเมินจะสามารถยอมรับการเปลี่ยนแปลงเพื่อแก้ไขให้ถูกต้อง แต่ทั้งนี้ต้องยึดหลักการความเป็นอิสระของผู้ประเมินที่จะสามารถสรุปผลการประเมินจากข้อค้นพบด้วยตนเอง

วิธีการตรวจสอบ

- มีการระบุถึงจุดแข็ง จุดอ่อนของการประเมิน
- มีมาตรฐานในการตรวจสอบผลการประเมิน
- มีการระบุว่ามีการตรวจสอบผลการประเมิน จากตัวผู้ประเมิน หรือผู้ที่เกี่ยวข้องกับการประเมิน

เกณฑ์การประเมิน

1. **ไม่มี**การรายงานการตรวจสอบผลการประเมินจากผู้ที่เกี่ยวข้องกับการประเมิน
2. มีการกำหนดแนวทางเพื่อตรวจสอบผลการประเมิน โดยผู้ที่มีส่วนเกี่ยวข้องกับการประเมิน
3. มีการตรวจสอบความถูกต้อง ของผลการประเมินตามมาตรฐานการประเมินที่กำหนดขึ้น
4. มีการตรวจสอบความถูกต้อง นำเชื่อถือของผลการประเมิน และประเมินจุดแข็ง จุดอ่อนของการประเมิน โดยคณะผู้ประเมินเอง หรือผู้ที่มีส่วนเกี่ยวข้องกับการประเมิน

มาตรฐานที่ 2 รรถประโยชน์ (Utility)

คุณลักษณะของงานประเมิน ที่เป็นประโยชน์ต่อผู้มีส่วนได้ส่วนเสีย และผู้ที่เกี่ยวข้อง มีการตัดสินใจคุณค่า และรายงานผลการประเมินอย่างชัดเจน มีการเผยแพร่ผลการประเมินทันความต้องการของผู้ใช้ผลการประเมิน และเป็นประโยชน์ในการวางแผน ปรับปรุงการดำเนินงานในทิศทางที่เสนอแนะ ประกอบด้วย 10 ตัวบ่งชี้ คือ

Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องในการประเมิน

คำอธิบาย ผู้มีส่วนได้ส่วนเสียในการประเมิน หมายถึง ผู้ที่เกี่ยวข้องกับการประเมินและได้รับผลกระทบจากการประเมิน ได้แก่ ผู้บริหาร อาจารย์ บุคลากร นักศึกษา ศิษย์เก่า ผู้ใช้บัณฑิต ผู้ปกครอง และชุมชน เป็นต้น

วิธีการตรวจสอบ

- มีการระบุถึงบุคคลที่มีส่วนได้ส่วนเสียในการประเมิน ได้แก่ นักศึกษา ศิษย์เก่า ผู้ปกครอง และสถานประกอบ ชุมชน ไว้ในรายงาน
- มีการอธิบายถึงการมีส่วนร่วมของผู้มีส่วนได้ส่วนเสียในการวางแผนการประเมิน

เกณฑ์การประเมิน

1. ไม่มีการระบุผู้มีส่วนได้ส่วนเสียในการประเมิน
 2. มีการระบุผู้มีส่วนได้ส่วนเสียที่มีส่วนร่วมในการประเมินเพียงกลุ่มเดียว
 3. มีการระบุผู้มีส่วนได้ส่วนเสียอย่างน้อย 2 กลุ่ม เข้ามามีส่วนร่วมในการประเมิน
 4. มีการระบุถึงผู้มีส่วนได้ส่วนเสีย ครบทุกกลุ่ม เข้ามามีส่วนร่วมและเกี่ยวข้องกับการประเมิน
- Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม

คำอธิบาย รายงานผลการประเมินมีการระบุถึงช่วงเวลา ระยะเวลาในการเก็บรวบรวมข้อมูลประเมิน โดยมีกำหนดการประเมินที่ระบุรายละเอียดการเก็บรวบรวมข้อมูล ผู้ให้ข้อมูล และมีการประสานงานกับผู้ให้ข้อมูล ในการจัดเตรียมข้อมูลให้สอดคล้องกับการประเมิน

วิธีการตรวจสอบ

- มีการเขียนกำหนดการประเมินไว้ โดยระบุระยะเวลา วัน เดือน ปีที่ประเมิน
- มีรายละเอียดถึงช่วงเวลา หรือระยะเวลาในการเก็บรวบรวมข้อมูลไว้

เกณฑ์การประเมิน

1. **ไม่มีการระบุ** เกี่ยวกับช่วงเวลา และระยะเวลา ในการเก็บรวบรวมข้อมูล
2. มีการระบุกำหนดการในการเก็บรวบรวมข้อมูลไว้ เป็นประมาณการระยะเวลา ในการเก็บรวบรวมข้อมูลเท่านั้น
3. มีการระบุ ช่วงเวลาเพียงกว้างๆ ในการเก็บรวบรวมข้อมูล แต่ไม่ระบุรายละเอียดของระยะเวลาในการเก็บรวบรวมข้อมูลไว้
4. มีการระบุ ช่วงเวลา ระยะเวลา และรายละเอียดในการเก็บรวบรวมข้อมูล อย่างชัดเจน เพื่อให้ผู้รับการประเมินสามารถจัดเตรียมข้อมูล ให้สอดคล้องกับการประเมิน

Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล โดยคำนึงถึงข้อจำกัดของทรัพยากรการประเมิน

คำอธิบาย เทคนิคการเก็บรวบรวมข้อมูล ได้แก่ การเลือกสุ่มตัวอย่าง เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล ชนิดของข้อมูล โดยวางแผนการประเมินเพื่อรวบรวมสารสนเทศที่จำเป็น ซึ่งพิจารณาจากทรัพยากรที่จำเป็น ผลกระทบต่อข้อค้นพบและข้อสรุปในการประเมิน รวมทั้งข้อจำกัดในการประเมิน

วิธีการตรวจสอบ

- มีการระบุถึงเทคนิค และกระบวนการที่ใช้ในการประเมิน
- มีการอธิบายถึงวิธีการจัดการประเมินตามทรัพยากรที่มีอยู่อย่างจำกัด
- พิจารณาผลการประเมินว่าสอดคล้องตามความสนใจ ของผู้รับผลการประเมิน

เกณฑ์การประเมิน

1. **ไม่มี**การวางแผนเกี่ยวกับการกำหนดเทคนิคในการประเมินไว้ล่วงหน้า
2. มีการวางแผนกำหนดเทคนิคในการประเมินไว้ล่วงหน้า
3. มีการวางแผนกำหนดเทคนิคการประเมินไว้ ตามทรัพยากรที่มีอย่างจำกัด
4. มีการวางแผนกำหนดเทคนิคการประเมินตามทรัพยากรที่มีอย่างจำกัด โดยคำนึงถึงผลกระทบต่อผลการประเมิน และข้อจำกัดในการประเมิน

Ut 2.4 การตรวจสอบการตัดสินผลการประเมินสอดคล้องกับสภาพจริง

คำอธิบาย สารสนเทศในรายงานมีความสมบูรณ์ โครงสร้างครบถ้วน นำเสนอชัดเจน โดยการพิจารณาถึงความเหมาะสมในการตีความผลการประเมินที่มีหลักฐานที่น่าเชื่อถือ เพียงพอต่อการตัดสินคุณค่า ตามมาตรฐานการประเมินที่กำหนด และสอดคล้องกับสภาพจริง

วิธีการตรวจสอบ

- มีการระบุถึงการตรวจสอบผลการประเมินก่อนการตัดสินผล
- มีการอธิบายถึงการตัดสินผลการประเมิน โดยพิจารณาถึงแหล่งข้อมูล และหลักฐาน ยืนยันตามความเป็นจริง

เกณฑ์การประเมิน

1. **ไม่มี** การระบุพิจารณาการตัดสินผลการประเมินตามมาตรฐานการประเมิน
2. มีการพิจารณาตัดสินผลการประเมิน ตามมาตรฐานการประเมิน แต่**ไม่มี**การอ้างอิงหลักฐาน
3. มีการพิจารณาการตัดสินผลการประเมินตามมาตรฐานการประเมิน ผลการวิเคราะห์และพยายามอ้างอิงหลักฐานประกอบ แต่ยังไม่ชัดเจน
4. มีการพิจารณาการตัดสินผลการประเมินที่ชัดเจนตามมาตรฐานการประเมิน ผลการวิเคราะห์ โดยอ้างอิงหลักฐานที่น่าเชื่อถือตามสภาพจริง

Ut 2.5 ผลการประเมินเป็นประโยชน์ ซึ่งให้เห็นทิศทางการปรับปรุง การพัฒนาการดำเนินงานของสถาบัน

คำอธิบาย ผลการประเมินและข้อสรุป สอดคล้องกับวัตถุประสงค์ของการประเมินและการตัดสินโดยอิงจากการประเมิน โดยต้องมีความชัดเจน และสามารถนำไปใช้ประโยชน์ในการวางแผนดำเนินการ ซึ่งให้เห็นทิศทางการพัฒนาที่ปฏิบัติได้จริงทันเหตุการณ์

วิธีการตรวจสอบ

- มีการระบุข้อเสนอแนะที่เป็นประโยชน์ และทิศทางการปรับปรุง พัฒนาการดำเนินงานของสถาบัน
- มีหลักฐานแสดงข้อมูลย้อนกลับจากกลุ่มผู้รับผลการประเมินเพื่อนำผลการประเมินไปใช้

เกณฑ์การประเมิน

1. ผลการประเมิน **ไม่มี** การระบุข้อเสนอแนะที่นำไปใช้ประโยชน์ และปรับปรุงการดำเนินงาน

2. ผลการประเมินเป็นประโยชน์ในการวางแผนการดำเนินงาน แต่ข้อเสนอแนะ ยังไม่สามารถนำไปปฏิบัติได้

3. ผลการประเมินเป็นประโยชน์ในการวางแผนการดำเนินงาน ปรับปรุงตามข้อเสนอแนะ แต่ไม่มีการประชุมเพื่อรับฟังข้อมูลย้อนกลับจากการใช้ผลการประเมิน

4. ผลการประเมินเป็นประโยชน์ในการวางแผน การดำเนินงาน โดยข้อเสนอแนะจากการประเมินชี้ให้เห็นทิศทางการพัฒนา มีการประชุมเพื่อรับฟังข้อมูลย้อนกลับจากการใช้ผลการประเมิน

Ut 2.6 การเขียนรายงานสอดคล้องกับรูปแบบที่กำหนด

คำอธิบาย รายงานผลการประเมินตามรูปแบบที่สำนักงานคณะกรรมการการอุดมศึกษา กำหนด ประกอบด้วย รายงานคณะผู้ประเมิน บทนำ การวางแผนและการประเมิน วิธีการตรวจสอบความถูกต้องน่าเชื่อถือของข้อมูล ผลการประเมินรายตัวบ่งชี้ องค์ประกอบคุณภาพมาตรฐานการอุดมศึกษา และมุมมองด้านการบริหารจัดการ มีจุดแข็ง จุดอ่อน และข้อเสนอแนะของแต่ละองค์ประกอบและมาตรฐาน รวมถึงภาคผนวก

วิธีการตรวจสอบ

- พิจารณาจากส่วนประกอบของรายงานถูกต้อง ครบถ้วนตามที่สกอ.กำหนด (ดูจากนิยามศัพท์ รูปแบบรายงาน สกอ.)
- พิจารณาความถูกต้องของการพิมพ์ การจัดเรียงหน้า

เกณฑ์การประเมิน

1. การเขียนรายงานยังไม่ครบถ้วน และไม่ถูกต้องตามรูปแบบที่สกอ.กำหนดไว้

2. การเขียนรายงานการประเมินเป็นไปตามรูปแบบที่กำหนด แต่ขาดยังส่วนสาระสำคัญบางส่วน ได้แก่ ส่วนสรุป หรือวิธีการประเมิน หรือข้อเสนอแนะ

3. การเขียนรายงานการประเมิน ครบถ้วนทั้งรายชื่อคณะผู้ประเมิน บทนำ วิธีการประเมิน ผลการประเมินและข้อเสนอแนะ ภาคผนวก ตามรูปแบบที่สกอ.กำหนด อาจมีการพิมพ์และจัดหน้าผิดพลาดเล็กน้อย

4. การเขียนรายงานการประเมินถูกต้อง ครบถ้วน ทั้งรายชื่อคณะผู้ประเมิน บทนำ วิธีการประเมิน ผลการประเมินและข้อเสนอแนะ ภาคผนวก ตามรูปแบบที่สกอ.กำหนด จัดพิมพ์ถูกต้อง

Ut 2.7 รายงานผลการประเมินครบถ้วน ครอบคลุมผลการประเมิน และข้อเสนอแนะเพื่อการพัฒนา

คำอธิบาย รายงานผลการประเมินมีการนำเสนอสารสนเทศที่สำคัญในรายงานการประเมินครบถ้วน ครอบคลุมผลการประเมิน โดยแสดงจุดแข็งและจุดอ่อนของข้อสรุปในการประเมิน ซึ่งข้อสรุปในการประเมินอิงจากข้อค้นพบ ที่ตอบคำถามในการประเมินตามข้อมูลที่สืบค้นได้ โดยระบุปัญหาและข้อเสนอแนะที่สามารถนำไปปฏิบัติได้ และปรับปรุงการประเมินครั้งต่อไป

วิธีการตรวจสอบ

- พิจารณาความครบถ้วน ของเนื้อหาในรายงานการประเมิน
- มีการระบุจุดเด่น จุดที่ควรพัฒนา และข้อเสนอแนะไว้ในรายงาน
- พิจารณาว่าข้อเสนอแนะ เป็นประโยชน์และสถาบันสามารถนำไปปฏิบัติได้อย่างรวดเร็ว

เกณฑ์การประเมิน

1. รายงานการประเมินยังมีเนื้อหาไม่ครบถ้วน ครอบคลุมผลการประเมิน
2. รายงานการประเมินมีเนื้อหาครบถ้วน ครอบคลุมผลการประเมิน จุดเด่น จุดที่ควรพัฒนา บางองค์ประกอบ แต่ขาดข้อเสนอแนะที่เป็นประโยชน์
3. รายงานการประเมินมีเนื้อหาครบถ้วนตามที่กำหนด ครอบคลุมผลการประเมิน มีการเสนอจุดเด่น จุดที่ควรพัฒนา แต่ไม่ระบุสาเหตุของปัญหา และข้อเสนอแนะ นำไปปฏิบัติได้ยาก
4. รายงานการประเมินมีเนื้อหาครบถ้วนตามที่กำหนด ครอบคลุมผลการประเมิน เสนอจุดเด่น จุดที่ควรพัฒนา ระบุสาเหตุของปัญหาและข้อเสนอแนะที่เป็นประโยชน์นำไปปฏิบัติได้

Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ และมีสื่อประกอบ

คำอธิบาย รายงานผลการประเมินจะต้องมีความชัดเจน เขียนด้วยภาษาเชิงวิชาการที่อ่านเข้าใจง่าย และบอกถึงข้อจำกัดของการประเมิน โดยมีสื่อประกอบด้วย รูปภาพ กราฟ ตาราง และตัวอย่างที่ง่ายต่อความเข้าใจ

วิธีการตรวจสอบ

- พิจารณาภาษาที่เขียนในรายงาน บุคคลทั่วไปอ่านแล้วเข้าใจง่าย
- มีสื่อประกอบ เช่น รูปภาพ กราฟ ตาราง ที่น่าสนใจ และเข้าใจได้ง่าย

เกณฑ์การประเมิน

1. การเขียนรายงานการประเมิน ใช้ภาษาที่ไม่ชัดเจน อ่านไม่เข้าใจ และไม่มีสื่อประกอบ

2. การเขียนรายงานการประเมิน ใช้ภาษาที่ค่อนข้างชัดเจน แต่ไม่มีสื่อ ประเภทตาราง ข้อมูลประกอบ

3. การเขียนรายงานการประเมิน ใช้ภาษาที่ชัดเจน เน้นข้อความตัวอักษร มีสื่อที่เป็นตารางประกอบบางส่วน

4. การเขียนรายงานการประเมิน ใช้ภาษาเชิงวิชาการที่ชัดเจน เข้าใจง่าย และมีการนำเสนอด้วยสื่อ เช่น รูปภาพ กราฟ ที่น่าสนใจประกอบ

Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา

คำอธิบาย รายงานผลการประเมินต้องมีความครบถ้วน สมบูรณ์ จัดทำรายงานเสร็จทันตามกำหนด และตรงกับช่วงเวลาที่ใช้ข้อมูลต้องการนำไปพิจารณาวางแผน ปรับปรุงการจัดการดำเนินการในองค์กร และการประเมิน เป็นไปตามกฎกระทรวงว่าด้วย ระบบ หลักเกณฑ์ และวิธีการประกันคุณภาพการศึกษาภายในสถานศึกษา ระดับอุดมศึกษา พ.ศ. 2546

วิธีการตรวจสอบ

- พิจารณารวันที่ส่งรายงาน เป็นไปตามกำหนดสกอ.หรือไม่ ประมาณ 30 กันยายน
- พิจารณาความสมบูรณ์ของรายงาน ว่าให้ข้อมูลที่ครบถ้วน เหมาะสมสอดคล้องของผลการประเมินที่เป็นประโยชน์ต่อผู้รับผลการประเมินในการนำไปพิจารณาปรับปรุงอย่างเร่งด่วน

เกณฑ์การประเมิน

1. การส่งรายงานการประเมินล่าช้า กว่ากำหนด(30 ก.ย.)
2. การจัดทำรายงานการประเมินมีข้อมูลที่สมบูรณ์ แต่ยังไม่เสร็จล่าช้า กว่ากำหนด (30 ก.ย.)
3. มีการจัดทำรายงานการประเมินเสร็จทันเวลาที่กำหนด แต่ยังไม่ขาดความสมบูรณ์ในบางส่วน
4. การจัดทำรายงานการประเมิน มีข้อมูลที่สมบูรณ์ เสร็จทันเวลาที่กำหนด ตรงช่วงเวลาที่ใช้ผลการประเมินต้องการ

Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม

คำอธิบาย มีการเผยแพร่ผลการประเมินในช่วงเวลาที่เหมาะสมตามความต้องการของผู้ใช้ผลการประเมิน และเปิดเผยผลการประเมินอย่างกว้างขวาง ใช้ภาษาเชิงวิชาการที่ครอบคลุมผู้ใช้ผลการประเมินทุกกลุ่ม

วิธีการตรวจสอบ

- พิจารณารายงานผลการประเมิน ว่ามีการนำเสนอผลการประเมินที่ทุกกลุ่มสามารถเข้าใจง่าย โดยใช้ภาษาที่ไม่ซับซ้อน
- พิจารณาว่าผลการประเมินนำเสนออย่างตรงไปตรงมา ครบถ้วน

เกณฑ์การประเมิน

1. มีการเผยแพร่ผลการประเมินใช้ภาษาพูด เน้นเฉพาะกลุ่มผู้รับการประเมิน
2. มีการเผยแพร่ผลการประเมินด้วยภาษาที่กำกวม สามารถเข้าใจได้เฉพาะกลุ่ม
3. มีการเผยแพร่ผลการประเมิน ด้วยภาษาเขียนที่สามารถเข้าใจได้ง่าย
4. มีการเผยแพร่ผลการประเมินอย่างเปิดเผย ด้วยภาษาที่เหมาะสมสื่อความได้ต่อความเข้าใจกับทุกคน

มาตรฐานที่ 3 จริยธรรมในการประเมิน (Ethicality)

คุณลักษณะของงานประเมินที่มีการกำหนดวิธีการประเมินไว้เหมาะสมกับสภาพจริง ได้รับการยอมรับ มีการคำนึงถึงผู้ที่เกี่ยวข้องหลายกลุ่ม มีการปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงสิทธิมนุษยชน และประโยชน์สาธารณะ ทำการประเมินที่มีความถูกต้องยุติธรรมต่อผู้รับผลการประเมิน รวมถึงเปิดเผยผลการประเมินแก่ผู้ที่ได้รับผลกระทบจากการประเมิน ประกอบด้วย 9 ตัวบ่งชี้ คือ

Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมินตลอดช่วงเวลาประเมิน

คำอธิบาย การจัดการประเมิน ที่มีคณะผู้ประเมินที่มีศักยภาพในการประเมิน สามารถกำหนดวิธีการดำเนินงานและตารางการปฏิบัติงานประเมินให้เหมาะสมตามสภาพจริง โดยมีการเจรจากับผู้ที่เกี่ยวข้องที่มีส่วนร่วมในการประเมินตลอดช่วงเวลาในการทำการประเมิน

วิธีการตรวจสอบ

- มีการระบุแผนการประเมิน วิธีการ กระบวนการประเมิน และการรวบรวมสารสนเทศในการประเมิน
- มีการอธิบายถึง การสื่อสาร ด้วยการซักถาม สืบค้นข้อมูล จากผู้รับผลการประเมิน ในระหว่างที่ทำการประเมิน

เกณฑ์การประเมิน

1. **ไม่มี**การสื่อสารวางแผนกำหนดวิธีการประเมินตามสภาพความเป็นจริง
2. มีการสื่อสารวางแผนกำหนดวิธีการประเมิน แต่ยังไม่สอดคล้องกับสภาพจริง ทำให้เป็นภาระของสถาบันในการประเมิน
3. มีการสื่อสารกำหนดวิธีการประเมิน คณะผู้ประเมิน ที่สอดคล้องกับสภาพความเป็นจริง แต่ **ไม่มี**หลักฐานที่ชัดเจนในการเจรจากับผู้มีส่วนร่วมในการประเมิน
4. มีการสื่อสารกำหนดวิธีการประเมินคณะผู้ประเมิน ที่เหมาะสมกับสภาพความเป็นจริง โดยมีการเจรจากับผู้มีส่วนร่วมในการประเมิน ตลอดช่วงระยะเวลาทำการประเมิน

Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง และปราศจากอคติการประเมิน

คำอธิบาย การประเมินโดยคำนึงถึงความแตกต่างของผู้ที่เกี่ยวข้องหลากหลายกลุ่ม ปราศจากอคติ และการตัดสินใจที่อาจขัดขวางการประเมิน ผลการประเมินเป็นที่ยอมรับจากสถาบัน โดยผู้ประเมินเสนอผลการประเมินตามวัตถุประสงค์ของการประเมิน ที่สะท้อนให้เห็นจุดเด่น/จุดที่ควรพัฒนา/ทิศทางการพัฒนาในอนาคต

วิธีการตรวจสอบ

- พิจารณาผลการประเมิน เป็นไปตามวัตถุประสงค์และเกณฑ์การประเมิน
- พิจารณาถึงข้อโต้แย้งที่เกี่ยวกับผลการประเมิน จากผู้ที่เกี่ยวข้อง
- พิจารณาถึงผลการประเมินที่สะท้อนให้เห็นจุดเด่น จุดที่ควรพัฒนา และทิศทางการพัฒนาของสถาบัน

เกณฑ์การประเมิน

1. ผลการประเมิน **ไม่**ได้รับการยอมรับจากบุคคลหลากหลายกลุ่ม
2. การประเมินบรรลุวัตถุประสงค์ เป็นไปตามเกณฑ์การประเมิน แต่มีการทักท้วงผลการประเมินเป็นลายลักษณ์อักษรจากผู้รับการประเมิน
3. การประเมินเป็นไปตามวัตถุประสงค์และเกณฑ์การประเมิน มีความพยายามประสานงานเพื่อเชื่อมโยงความเข้าใจกับบุคคลหลากหลายกลุ่ม แต่มีการทักท้วงผลการประเมินด้วยวาจา
4. ผลการประเมินเป็นที่ยอมรับจากบุคคลหลากหลายกลุ่ม ปราศจากอคติ ที่ขัดขวางการประเมิน ซึ่งผลการประเมินสะท้อนให้เห็นจุดเด่น จุดที่ควรพัฒนา และทิศทางการพัฒนาในอนาคต

Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงความต้องการจำเป็นของผู้รับการประเมิน

คำอธิบาย การประเมินคุณภาพ มีการพิจารณานำผลการประเมินมาปรับปรุงคุณภาพการประเมิน โดยเปรียบเทียบผลการประเมินกับความต้องการจำเป็นของผู้รับผลการประเมิน เพื่อให้ได้รับประโยชน์สูงสุดจากการประเมินอย่างต่อเนื่อง

วิธีการตรวจสอบ

- มีการระบุถึงการนำผลการประเมินที่ผ่านมา เพื่อปรับปรุงคุณภาพการประเมิน
- มีการอธิบายถึงการเปรียบเทียบผลการประเมินกับความต้องการของผู้รับผลการประเมิน

เกณฑ์การประเมิน

1. **ไม่**มีการทบทวนแผนการประเมิน และพิจารณาปรับปรุงคุณภาพผลการประเมิน
2. มีการทบทวนวางแผนการประเมินและพิจารณาปรับปรุงคุณภาพผลการประเมิน แต่มีรายละเอียดไม่ชัดเจน

3. มีการพิจารณาปรับปรุงคุณภาพผลการประเมิน ระหว่างการประเมิน แต่ยังไม่ครอบคลุมความต้องการจำเป็นของผู้รับการประเมิน

4. มีการพิจารณาปรับปรุงคุณภาพผลการประเมิน ระหว่างการประเมิน และอธิบายเปรียบเทียบผลการประเมินกับความต้องการของผู้รับการประเมิน

Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน

คำอธิบาย มีการทำเอกสารการประเมิน กำหนดการประเมินและวิธีการประเมินที่ประยุกต์ใช้เฉพาะกลุ่มที่รับการประเมินและผู้มีส่วนได้ส่วนเสียอื่นที่เกี่ยวข้องอย่างเป็นลายลักษณ์อักษรระหว่างผู้ประเมินกับสถาบัน ตามวัตถุประสงค์การประเมิน กระบวนการประเมิน ทรัพยากรในการประเมิน

วิธีการตรวจสอบ

- มีเอกสารที่ระบุถึงการตกลงทำการประเมิน ระหว่างสถาบัน ได้แก่ กำหนดการประเมิน วิธีการประเมิน
- มีการอธิบายถึงแหล่งทรัพยากรที่ใช้ในการประเมิน

เกณฑ์การประเมิน

1. มีการจัดทำกำหนดการประเมินระหว่างผู้ประเมินกับสถาบัน
2. มีการระบุกำหนดการประเมิน และวิธีการประเมินระหว่างผู้ประเมินกับสถาบัน
3. มีการตรวจสอบ ทบทวน กำหนดการและวิธีการประเมิน ที่สอดคล้องกับทรัพยากรในการประเมิน
4. มีการดำเนินการประเมิน ตามกำหนดการและวิธีการประเมิน ที่สอดคล้องกับวัตถุประสงค์ กระบวนการ และทรัพยากรในการประเมิน

Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัดที่มีผลกระทบต่อผลการประเมิน

คำอธิบาย การประเมินมีการเปิดเผยผลการประเมิน ข้อจำกัดในการประเมินให้กลุ่มบุคคลทุกฝ่ายที่ได้รับผลกระทบ หรือบุคคลอื่นที่เกี่ยวข้องได้รับทราบผลการประเมิน โดยรายงานผลการประเมินมีรายละเอียดที่ครอบคลุม ตอบสนองความต้องการของผู้รับผลการประเมิน

วิธีการตรวจสอบ

- มีการอธิบายถึงการเปิดเผยผลการประเมินให้ผู้ที่เกี่ยวข้องทราบ
- มีการอธิบายรายละเอียดข้อจำกัดของผลการประเมินให้ผู้ที่เกี่ยวข้องทุกกลุ่มทราบ

เกณฑ์การประเมิน

1. **ไม่มีการ**เปิดเผยผลการประเมินและข้อจำกัดที่ส่งผลกระทบต่อผลการประเมินให้ผู้ที่เกี่ยวข้องรับทราบ

2. ผู้ที่เกี่ยวข้องรับทราบเฉพาะผลการประเมิน แต่ **ไม่มีการ**เปิดเผยข้อจำกัดที่ส่งผลกระทบต่อ การประเมิน

3. ผู้ที่เกี่ยวข้องรับทราบผลการประเมินและข้อจำกัดที่ส่งผลกระทบต่อ การประเมิน แต่ยังมี รายละเอียดไม่ชัดเจน ครอบคลุมทุกองค์ประกอบ

4. ผู้ที่เกี่ยวข้องรับทราบผลการประเมินและข้อจำกัดที่ส่งผลกระทบต่อ การประเมิน อย่าง เปิดเผย ชัดเจน ครอบคลุมทุกองค์ประกอบ

Et 3.6 การกำหนดวิธีการประเมินโดยคำนึงถึงความปลอดภัย รักษาความลับ และไม่ ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล

คำอธิบาย การกำหนดขอบเขตและวิธีการประเมินการประเมินที่คำนึงถึงความเท่าเทียม คุณค่า ความปลอดภัย รักษาความลับ และพิจารณาสิทธิมนุษยชนของผู้รับการประเมิน และผู้มีส่วนได้ส่วนเสียอื่นๆ ในการออกแบบและความแตกต่างของสารสนเทศในการประเมิน โดยไม่ ก่อให้เกิดความเสียหายและความทุกข์ใจที่จะเกิดกับผู้มีส่วนร่วมในการประเมิน

วิธีการตรวจสอบ

- มีการระบุคำพูดของบุคคลโดยบอกชื่อหรือตำแหน่ง ของผู้ที่เกี่ยวข้องกับการประเมิน ที่สร้างความเสียหายหรืออันตราย
- มีการหลักฐานที่บ่งบอกข้อมูลที่เป็นความลับของผู้ที่เกี่ยวข้องกับการประเมิน

เกณฑ์การประเมิน

1. มีการเปิดเผยข้อมูลความลับ ที่ก่อให้เกิดความเสียหายและอันตรายแก่ผู้มีส่วนร่วมในการ ประเมิน

2. มีวิธีการรักษาข้อมูลความลับ โดยไม่ระบุตัวตน ของผู้มีส่วนร่วมในการประเมิน กรณีที่อาจ สร้างความเสียหายหรืออันตรายได้

3. มีวิธีการรักษาข้อมูลความลับ เช่น ไม่ระบุคำพูด ของผู้มีส่วนร่วมในการประเมิน กรณีที่ สามารถสร้างความเสียหายหรืออันตรายได้

4. **ไม่มีการ**เปิดเผย ข้อมูลความลับของผู้มีส่วนร่วมในการประเมิน กรณีที่สร้างความเสียหาย หรืออันตรายโดยตรงได้

Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง โดยคำนึงถึงความแตกต่าง หลากหลายของผู้ที่เกี่ยวข้อง

คำอธิบาย การจัดการประเมินที่คำนึงถึงความสัมพันธ์กับผู้ที่เกี่ยวข้องในการประเมินที่มี ความแตกต่าง และหลากหลาย โดยมีการศึกษาถึงภูมิหลังของบุคคลที่เกี่ยวข้องกับการประเมิน เพื่อสร้างปฏิสัมพันธ์ที่ดี

วิธีการตรวจสอบ

- มีการอธิบายถึงการศึกษานิวทอนที่แตกต่างกันของผู้ที่เกี่ยวข้องกับการประเมิน
- มีการอธิบายถึงการสื่อสารทั้งวาจา เอกสาร หรือช่องทางอื่น เพื่อแสดงความสัมพันธ์กับผู้ที่เกี่ยวข้องกับการประเมิน

เกณฑ์การประเมิน

1. **ไม่มี** ข้อมูลนิวทอนของผู้ที่เกี่ยวข้องกับการประเมิน ได้แก่ นักศึกษา ศิษย์เก่า ผู้ปกครอง อาจารย์ และผู้เข้าบัณฑิต เป็นต้น
2. มีการสอบถามถึงนิวทอนของผู้ที่เกี่ยวข้องกับการประเมิน แต่ยังไม่มีการสื่อสารและสร้างความสัมพันธ์กับผู้ที่เกี่ยวข้องกับการประเมิน
3. มีการศึกษานิวทอนของผู้ที่เกี่ยวข้องกับการประเมิน และพยายามสร้างความสัมพันธ์กับผู้ที่เกี่ยวข้องกับการประเมินที่มีพื้นฐานแตกต่างกัน
4. มีการตรวจสอบนิวทอนทางของผู้ที่เกี่ยวข้องกับการประเมิน และดำเนินการประเมินโดยมีการสื่อสารข้อมูลและสร้างความสัมพันธ์กับผู้ที่เกี่ยวข้องในการประเมินพื้นฐานแตกต่างกัน

Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย

คำอธิบาย การดำเนินการประเมินที่ให้ความเป็นธรรมกับผู้ที่เกี่ยวข้องกับการประเมิน โดยมีการตรวจสอบและบันทึกจุดเด่น จุดที่ควรพัฒนาของสถาบันที่ทำการประเมินเพื่อให้ข้อเสนอแนะที่มีความเป็นกลางต่อผู้ที่นำผลการประเมินไปใช้ประโยชน์

วิธีการตรวจสอบ

- มีการระบุถึงการตรวจสอบความถูกต้องของฐานข้อมูลในการประเมิน
- จุดเด่น จุดที่ควรพัฒนามีความเป็นธรรมต่อผู้ใช้ผลการประเมิน
- ข้อเสนอแนะมีความเป็นธรรมต่อผู้ใช้ผลการประเมิน

เกณฑ์การประเมิน

1. **ไม่มี** การตรวจสอบความถูกต้องของการตัดสินผลการประเมิน
2. มีการตรวจสอบความถูกต้องของฐานข้อมูลก่อนการตัดสินผลการประเมิน แต่ไม่ได้พิจารณาข้อเสนอแนะที่เป็นกลางต่อผู้ใช้ผลการประเมิน
3. มีการตรวจสอบความถูกต้องของฐานข้อมูลก่อนการตัดสินผลการประเมิน มีความพยายามพิจารณาข้อเสนอแนะที่เป็นกลางต่อผู้ใช้ผลการประเมิน
4. มีการตัดสินผลการประเมินจากฐานข้อมูลที่ถูกต้อง ครบถ้วน ตรวจสอบจุดเด่น จุดที่ควรพัฒนา โดยให้ข้อเสนอแนะที่เป็นกลางกับผู้ที่เกี่ยวข้องกับการประเมินไปใช้

Et 3.9 การประเมินที่ครบถ้วน ชัดเจนสอดคล้องกับมาตรฐานและ ตัวบ่งชี้

คำอธิบาย การตัดสินผลการประเมินต้องมาจากฐานข้อมูลที่ถูกต้อง ครบถ้วน ทั้งการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การลงข้อสรุปข้อมูล การรายงานผลการประเมินสอดคล้องกับมาตรฐานในการประเมิน

วิธีการตรวจสอบ

- มีการระบุถึงการตรวจสอบความถูกต้องของข้อมูล การวิเคราะห์ข้อมูล การสรุปและนำไปใช้ประโยชน์
- มีการรวบรวมข้อมูล วิเคราะห์ข้อมูลและสรุปผลการประเมินสอดคล้องกับมาตรฐาน
- มีข้อเสนอแนะที่เป็นประโยชน์และสอดคล้องกับมาตรฐาน

เกณฑ์การประเมิน

1. **ไม่มีการ**ตรวจสอบความถูกต้องของข้อมูล การวิเคราะห์ข้อมูล และความสอดคล้องของการสรุปข้อมูล นำไปใช้ประโยชน์
2. มีการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล สอดคล้องกับการสรุปข้อมูล นำไปใช้ประโยชน์ในบางองค์ประกอบ
3. มีการเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล และสรุปข้อมูลรายงานผลการประเมินถูกต้อง ครบถ้วน สอดคล้องกับองค์ประกอบแต่ข้อเสนอแนะยังไม่ชัดเจน นำไปปฏิบัติได้ยาก
4. มีการเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล และสรุปข้อมูลรายงานผลการประเมินถูกต้อง ครบถ้วน สอดคล้องกันทุกองค์ประกอบและมีข้อเสนอแนะที่เป็นประโยชน์และสอดคล้องกันตลอด

มาตรฐานที่ 4 ความน่าเชื่อถือของนักประเมิน (Credibility)

คุณลักษณะของงานประเมินที่ได้มาจากนักประเมินที่มีสมรรถนะในการประเมิน ได้รับการยอมรับ และไม่มีผลประโยชน์ทับซ้อนเพื่อสร้างความน่าเชื่อถือของข้อค้นพบและสารสนเทศที่ได้รับจากการประเมิน ประกอบด้วย 4 ตัวบ่งชี้ คือ

Cr4.1 ผู้ประเมินมีความสามารถ ในการประเมิน

คำอธิบาย ผู้ประเมินต้องมีสมรรถนะในการประเมิน ได้แก่ มีความรู้ ความสามารถ ทักษะ และความชำนาญหรือประสบการณ์ในการประเมินที่เหมาะสม โดยสามารถวัด วิเคราะห์ ข้อมูล แปลผล และเผยแพร่ข้อค้นพบอย่างชัดเจน มีเหตุผลเพียงพอสำหรับการตัดสินผลการประเมิน เขียนรายงานผลการประเมินได้ชัดเจน นอกจากนี้ใช้เทคนิคการดำเนินการทั้งเชิงคุณภาพและเชิงปริมาณ

วิธีการตรวจสอบ

- การวิเคราะห์ผลการประเมินมีความชัดเจน สอดคล้องกับฐานข้อมูลตลอดเล่ม

- ผลการประเมินมีความโปร่งใส ยุติธรรม สามารถตรวจสอบได้ และสอดคล้องกับมาตรฐานการประเมิน

เกณฑ์การประเมิน

1. ผู้ประเมินวิเคราะห์ ผลการประเมิน และเขียนรายงานขัดแย้งกัน ตลอดเล่ม
2. ผู้ประเมินวิเคราะห์ผลการประเมินและเสนอรายงานผลการประเมินไม่ชัดเจน การสรุปผลไม่มีฐานข้อมูลรองรับ
3. ผู้ประเมินวิเคราะห์ผลการประเมิน และเสนอรายงานผลการประเมินอย่างชัดเจน สอดคล้องกับสรุปผลและมีฐานข้อมูลรองรับ บางส่วน
4. ผู้ประเมินวิเคราะห์ผลการประเมิน และเสนอรายงานผลการประเมินอย่างโปร่งใส ละเอียดชัดเจน สอดคล้องกับสรุปผล และมีฐานข้อมูลรองรับ ส่วนใหญ่

Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน

คำอธิบาย ผู้ประเมินมีทักษะระหว่างบุคคล ได้แก่ ทักษะการสื่อสารทางการเขียน การนำเสนอผลงาน และทักษะต่างวัฒนธรรม ทักษะแก้ปัญหาความขัดแย้ง ส่งเสริมให้เกิดปฏิสัมพันธ์ระหว่างบุคคล มีการบันทึกจุดแข็งและข้อจำกัดของการประเมิน รวมถึงทักษะการนำเสนอข้อเสนอแนะเพื่อพัฒนาที่สามารถนำไปปฏิบัติได้

วิธีการตรวจสอบ

- มีการเสนอแนวทางเสริม และปรับปรุงจุดอ่อน ที่เป็นประโยชน์
- มีการอธิบายแนวทางการประเมินที่ชัดเจน โดยให้ทิศทางการวางแผนและพัฒนาอย่างเหมาะสม ที่สามารถปฏิบัติได้

เกณฑ์การประเมิน

1. ผู้ประเมินไม่สามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน ที่เป็นประโยชน์ได้
2. ผู้ประเมินสามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน แต่ไม่สามารถปฏิบัติเป็นรูปธรรมได้
3. ผู้ประเมินสามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน ครอบคลุม ผลการประเมิน แต่ยังมีรายละเอียดกว้างๆ นำไปปฏิบัติได้บางส่วน
4. ผู้ประเมินสามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน ที่ชัดเจนทำให้เห็นทิศทางการวางแผน และพัฒนาอย่างเหมาะสม นำไปปฏิบัติได้

Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมินอย่างเหมาะสม

คำอธิบาย ผู้ประเมินมีความสามารถในการจัดการประเมินตั้งแต่เริ่มต้นจนกระทั่งเสร็จสิ้น โดยสามารถสร้างกรอบแนวทางการประเมินที่สอดคล้องกับบริบทของสถาบัน มีการออกแบบการประเมินอย่างรัดกุม โดยประสานงานเกี่ยวกับงบประมาณ การใช้ทรัพยากรร่วมกันเช่น ข้อมูล ความ

เชี่ยวชาญ บุคลากร เครื่องมือปฏิบัติงาน มีการใช้เทคโนโลยีให้เหมาะสม และจัดการประเมินให้เหมาะสมกับกับช่วงระยะเวลาที่กำหนด มีการเสนอทิศทางการพัฒนาสถาบันตามจุดเน้น เพื่อให้บรรลุเป้าหมาย

วิธีการตรวจสอบ

- มีการเสนอแนวทางการพัฒนาที่สอดคล้องกับบริบทของสถาบัน
- มีการอธิบายถึงแนวทางการพัฒนาเป็นไปตามจุดเน้นของสถาบัน และสามารถบรรลุเป้าหมายได้

เกณฑ์การประเมิน

1. ผู้ประเมินยังไม่สามารถเข้าใจบริบท และมีมุมมองในการบริหารจัดการ และแนวทางการพัฒนาที่แคบ ยากต่อการพัฒนาสถาบัน
2. ผู้ประเมินเข้าใจบริบทของสถาบันไม่ชัดเจน ทำให้เสนอแนวทางการพัฒนาไม่สอดคล้องกับบริบทของสถาบัน
3. ผู้ประเมินสามารถเข้าใจบริบทของสถาบัน และมีวิธีการที่จะเสนอแนวทางการพัฒนาอย่างเหมาะสมในภาวการณ์ปัจจุบัน
4. ผู้ประเมินสามารถเข้าใจบริบทของสถาบัน และเสนอ ทิศทางการพัฒนาสถาบันตามจุดเน้น จนสามารถบรรลุเป้าหมายสูงสุดได้

Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อนของผู้ประเมิน

คำอธิบาย ผู้ประเมินต้องคำนึงถึงผลประโยชน์ที่เกี่ยวข้องกับสาธารณชนและสวัสดิภาพของสังคมเป็นหลัก โดยมีการเปิดเผยผลประโยชน์ทับซ้อนที่เกิดขึ้นอย่างซื่อสัตย์ เพื่อไม่ให้เกิดการประนีประนอมหรือละเว้นในกระบวนการประเมินและผลการประเมินที่มีอิทธิพลจากกลุ่มบุคคล ที่มีความขัดแย้งระหว่างกลุ่มที่เกี่ยวข้องกับการประเมิน

วิธีการตรวจสอบ

- มีหลักฐานหรือการยืนยันผลการประเมินที่ปราศจากความขัดแย้ง
- มีหลักฐานระบุถึงผลประโยชน์ทับซ้อนของผู้ประเมิน
- มีการอธิบายถึงการจัดการความขัดแย้ง ระหว่างกลุ่มที่เกี่ยวข้องกับการประเมิน

เกณฑ์การประเมิน

1. การตัดสินผลการประเมินมีความคลาดเคลื่อน เนื่องจากได้รับอิทธิพลเกี่ยวกับผลประโยชน์ที่ขัดแย้งกับผู้ประเมิน
2. การตัดสินผลการประเมินมีความชัดเจน แต่มีอิทธิพลเกี่ยวกับผลประโยชน์จากผู้ประเมินที่ส่งผลต่อการประเมิน

3. การตัดสินใจผลการประเมินมีความถูกต้อง ชัดเจน และ**ไม่ได้รับอิทธิพล**เกี่ยวกับผลประโยชน์จากผู้ประเมินที่ส่งผลต่อการประเมิน

4. การตัดสินใจผลการประเมินโปร่งใส ตรวจสอบได้ **ไม่มีอิทธิพล**เกี่ยวกับผลประโยชน์จากผู้ประเมินที่ส่งผลต่อการประเมิน

มาตรฐานที่ 5 ประสิทธิภาพต้นทุน (Cost-effectiveness)

คุณลักษณะของงานประเมิน ที่คำนึงถึงความคุ้มค่าของทรัพยากรที่จำเป็นในการประเมิน มีค่าใช้จ่ายในการประเมินอย่างสมเหตุสมผล มีหลักฐานชัดเจนและมีความโปร่งใสตรวจสอบได้ ของต้นทุนการประเมิน ประกอบด้วย 2 ตัวบ่งชี้ คือ

Ce5.1 การจัดการทรัพยากรในการประเมินอย่างคุ้มค่า

คำอธิบาย การประเมินที่ให้สารสนเทศที่มีคุณค่าเพียงพอสำหรับตัดสินใจผลการประเมิน สามารถประเมินได้บรรลุเป้าหมาย โดยมีการจัดการทรัพยากรในการประเมิน พิจารณาจากการใช้ทรัพยากรอย่างคุ้มค่า ได้แก่ ค่าใช้จ่ายที่เป็นเงิน เวลา บุคลากร สถานที่ในการจัดการประเมิน โดยผู้รับการประเมินมีความพึงพอใจ

วิธีการตรวจสอบ

- มีแผนการจัดการทรัพยากรการประเมิน
- มีการอธิบายถึงความชัดเจนของการตัดสินใจผลการประเมินจากข้อมูลสารสนเทศ
- มีการอธิบายถึงการประเมินที่คุ้มค่า คุ้มเวลา และคุ้มแรงงาน

เกณฑ์การประเมิน

1. **ไม่มีการ**วางแผนการจัดการทรัพยากรการประเมิน
2. มีการวางแผนการจัดการทรัพยากรการประเมิน แต่ไม่ระบุนการใช้ทรัพยากร
3. การจัดการทรัพยากรการประเมินเป็นไปตามเป้าหมายที่วางไว้อย่างเหมาะสม
4. การจัดการทรัพยากรการประเมิน อย่างประหยัดคุ้มค่าบรรลุเป้าหมาย และผู้รับการประเมินมีความพึงพอใจ

Ce5.2 การจัดสรรงบประมาณสำหรับการประเมินอย่างเหมาะสม

คำอธิบาย การจัดสรรงบประมาณเพื่อเป็นค่าใช้จ่ายในการประเมินอย่างสมเหตุสมผล มีหลักฐานแสดงชัดเจน ได้แก่ มีจัดทำโครงการ หรือจัดตั้งงบประมาณสำหรับการประเมินร่วมกับ สถาบันที่รับการประเมิน และระบุนายละเอียดค่าใช้จ่ายในการประเมินสามารถตรวจสอบได้ โดยมีการแสดงเหตุผลเกี่ยวกับต้นทุน และข้อมูลที่เป็นในการประเมิน

วิธีการตรวจสอบ

- มีการระบุโครงการ แผนงบประมาณในการประเมิน
- มีหลักฐานแสดงค่าใช้จ่ายที่ใช้ในการประเมิน ที่สามารถตรวจสอบได้

เกณฑ์การประเมิน

1. ไม่มีการวางแผนเกี่ยวกับงบประมาณในการประเมิน
2. มีค่าใช้จ่ายในการประเมิน แต่ไม่มีหลักฐานแสดงการใช้จ่าย ในการประเมิน
3. มีการสรุปค่าใช้จ่ายที่ใช้ในการประเมิน มีหลักฐานแสดงการใช้จ่ายในบางส่วน แต่ยังไม่ชัดเจน
4. มีเอกสารแสดงถึงค่าใช้จ่าย ในการประเมินอย่างสมเหตุสมผล จัดทำหลักฐานชัดเจน โปร่งใส สามารถตรวจสอบค่าใช้จ่าย ในการประเมินได้

4. คำนิยามศัพท์

4.1 **การวิเคราะห์ความแตกต่างของเอกสาร** หมายถึง การวิเคราะห์ความถูกต้อง น่าเชื่อถือ ของเอกสารที่รวบรวมมาจากแหล่งต่างกัน หรือบุคคลต่างกลุ่มกัน เพื่อพิจารณาความ สอดคล้องกันของข้อมูล ทั้งความครบถ้วนสมบูรณ์ ความน่าเชื่อถือ ความเป็นจริง และความถูกต้อง ในวิธีการวิเคราะห์ เสนอผลการประเมิน

4.2 **ข้อมูลพื้นฐานของสถาบัน** หมายถึง ข้อมูลที่เกี่ยวกับปรัชญา วิสัยทัศน์ โครงสร้าง ของสถาบัน จำนวนอาจารย์ บุคลากรนักศึกษา หลักสูตรที่เปิดสอน งบประมาณ เป็นต้น

4.3 **ภูมิหลังของสถาบัน** หมายถึง ประวัติความเป็นมาของสถาบัน การเปลี่ยนแปลง ตั้งแต่ก่อตั้ง จนถึงปัจจุบัน

4.4 **บริบทของสถาบัน** หมายถึง สถานที่ตั้ง องค์ประกอบภายในสถาบัน สภาพแวดล้อม โดยรวมของสถาบัน

4.5 **อัตลักษณ์ของสถาบัน** หมายถึง จุดเด่นของสถาบันที่แสดงถึงคุณลักษณะเฉพาะ ของสถาบันที่เป็นที่รู้จักของบุคคลโดยทั่วไป รวมเป้าหมายสูงสุดที่สถาบันต้องการจะเป็นใน อนาคต

4.6 **การออกแบบประเมินที่มีมุมมองต่างกัน** หมายถึง การวางแผนการประเมินที่ ครอบคลุมการประเมินกลุ่มผู้ประเมิน ผู้ใช้ผลการประเมิน และผู้ที่มีส่วนเกี่ยวข้องที่แตกต่างกัน

4.7 **การมีส่วนร่วมของผู้มีส่วนได้ส่วนเสีย** หมายถึง ผู้ที่มีส่วนได้ส่วนเสีย ได้แก่ ผู้บริหาร อาจารย์/บุคลากร นักศึกษา ศิษย์เก่า ผู้ปกครอง ผู้ใช้บัณฑิต และชุมชน เป็นต้น เข้ามา มีส่วนร่วมในการประเมิน ตั้งแต่ ให้ความร่วมมือ อำนวยความสะดวกในการประเมิน ให้ข้อมูล ประสานงาน และร่วมดำเนินการจัดการประเมิน

4.8 **ทรัพยากรที่มีจำกัด** หมายถึง ทรัพยากรเกี่ยวกับเงิน เวลา แรงงาน สถานที่ ที่มีอยู่ อย่างจำกัด เพื่อสามารถใช้ให้เกิดประโยชน์อย่างสูงสุด

4.9 แหล่งข้อมูลที่หลากหลาย หมายถึง ข้อมูลจากหลายแหล่ง ได้แก่ หลักฐานเอกสารอ้างอิง การสัมภาษณ์บุคคลากรที่เกี่ยวข้อง การสังเกตกิจกรรม สถานที่ รวมทั้งการสืบค้นจากฐานข้อมูลย้อนหลัง

4.10 รูปแบบรายงานของ สกอ. หมายถึง รายงานการประเมินคุณภาพที่ประกอบด้วย 1) คำนำ สารบัญ 2) ข้อมูลเบื้องต้นของสถาบัน ได้แก่ ประวัติ ปรัชญา ปณิธาน โครงสร้างการบริหารจัดการ ระบบและกลไกการประกันคุณภาพของสถาบัน สรุปผลการปรับปรุงตามผลการประเมินในรอบปีที่ผ่านมาและเป้าหมายสำคัญในปีปัจจุบัน 3) ผลการดำเนินงานและผลการประเมินคุณภาพเทียบกับเกณฑ์และเป้าหมายในรอบปีปัจจุบันตามตัวบ่งชี้ในแต่ละองค์ประกอบ 4) สรุปผลการประเมินและทิศทางการพัฒนา โดยแยกเป็นผลการประเมินในภาพรวมของแต่ละองค์ประกอบคุณภาพ และแต่ละมาตรฐานการอุดมศึกษา ตลอดจนจุดเด่น และจุดที่ควรปรับปรุงแผนและเป้าหมายการพัฒนาในปีต่อไป

4.11 การสื่อสารกับผู้รับผลการประเมิน หมายถึง การรายงานด้วยเอกสาร การสื่อสารทางวาจา และรับฟังคำอธิบาย เหตุผลของผู้รับผลการประเมิน ตั้งแต่การแจ้งผลการประเมินด้วยวาจา การเขียนรายงานที่ชัดเจน รวมถึงการเสนอแนะแนวทางการพัฒนาอย่างเป็นรูปธรรม

4.12 เอกสารข้อตกลงการประเมิน หมายถึง เอกสารที่แสดงถึงการดำเนินการจัดการประเมินคุณภาพภายใน ระดับสถาบัน โดยมีรายชื่อคณะผู้ประเมินที่มาจากภายนอกสถาบัน และภายในสถาบัน ตามที่สกอ.กำหนด รวมถึงวัตถุประสงค์ กระบวนการประเมิน และกำหนดการในการตรวจประเมินคุณภาพภายใน

4.13 ภูมิหลังและวัฒนธรรมทางสังคมของผู้ที่เกี่ยวข้องกับประเมิน หมายถึง ประวัติพอสังเขป ตำแหน่งหน้าที่ในปัจจุบัน ลำดับและความสำคัญต่อองค์กรของผู้ที่เกี่ยวข้องกับการประเมิน

4.14 ข้อจำกัดที่ส่งผลกระทบต่อผลการประเมิน หมายถึง ข้อขัดข้องต่างๆ ที่ส่งผลต่อการประเมิน ได้แก่ ช่วงเวลา และระยะเวลาในการประเมิน ความสมบูรณ์ เพียงพอของการรวบรวมข้อมูลเพื่อตัดสินผลการประเมิน ประสบการณ์ของผู้ประเมิน เป็นต้น

4.15 กลุ่มบุคคลที่ขัดแย้ง หมายถึง บุคคลที่อาจเสียผลประโยชน์ เมื่อมีการประเมินเกิดขึ้น โดยผู้ประเมินมีวิธีการที่สามารถบริหารจัดการเมื่อเกิดความขัดแย้ง ในระหว่างกระบวนการประเมิน

4.16 ต้นทุนทางการประเมิน หมายถึง ต้นทุนที่เกี่ยวข้องกับ เงิน เวลา แรงงาน ค่าเลือกราคาต่างๆ ในการทำการประเมิน

5. ลักษณะรายงานการประเมินที่ดี

รายงานการประเมินที่ดีควรมีลักษณะ 2 ประการ (สุชาติดา บวรกิติวงศ์, 2550) คือ

1. มีรูปแบบตรงตามที่สกอ. กำหนดประกอบด้วย 3 ส่วนหลัก คือ ส่วนนำ ได้แก่ ปก บทสรุป ผู้บริหาร และสารบัญ ส่วนเนื้อหา ได้แก่ สภาพทั่วไปของสถาบัน ผลการประเมินคุณภาพภายใน สรุปผลการประเมินคุณภาพภายในและส่วนภาคผนวก

2. มีเนื้อหาที่ครอบคลุม น่าเชื่อถือ เพียงพอที่จะทำให้ผู้ประเมินสามารถสรุปผลการประเมิน โดยคุณภาพของรายงานการประเมิน ควรประกอบด้วย 5 ประเด็น ดังต่อไปนี้ คือ

1. มีข้อเท็จจริงที่ถูกต้อง ควรใช้เทคนิคการตรวจสอบสามเส้า ได้แก่ เก็บรวบรวมข้อมูลจากหลายแหล่งเอกสาร เช่น นักศึกษา อาจารย์ ผู้บริหาร ผู้ใช้บัณฑิตและผู้ที่เกี่ยวข้อง เก็บข้อมูลจากหลากหลายวิธี เช่น การศึกษาเอกสาร การสังเกต การสัมภาษณ์ เก็บข้อมูลจากหลายสถานการณ์

2. มีผลการประเมินที่ถูกต้อง โดยตรวจสอบว่าการตัดสินแต่ละตัวบ่งชี้ว่าเป็นไปตามเกณฑ์ของสกอ. หรือไม่ รูปแบบการเขียนรายงานและการตรวจสอบผลการตัดสินว่าถูกต้องหรือไม่

3. ข้อสรุปสอดคล้องกับผลการประเมิน โดยเขียนรายงานสรุปไม่ควรขัดแย้งกับผลการประเมิน ได้แก่ ควรเสนอแนะกรณีที่ได้ผลการประเมินต่ำกว่าระดับดี และเขียนจุดเด่น กรณีมีผลการประเมินตั้งแต่ระดับดีขึ้นไป เป็นต้น

4. ข้อเสนอแนะสอดคล้องกับผลการประเมินของสถาบันที่ทำการประเมิน ข้อเสนอแนะที่ควรตรวจสอบ คือ 4.1) มีความสอดคล้อง มีความเป็นไปตามผลการประเมิน 4.2) ภายในมาตรฐานเดียวกันไม่ขัดแย้งกัน 4.3) ข้อเสนอแนะควรเป็นรูปธรรม ควรวางแนวทางให้ปฏิบัติได้และตอบสนองนโยบายสถาบัน หากเป็นไปได้ควรมีตัวอย่างชี้แนวทางปฏิบัติ

5. ข้อเสนอแนะเป็นประโยชน์สามารถนำไปปฏิบัติได้ โดยเป็นรูปธรรมให้สถาบันนำไปปฏิบัติได้จริง โดยคำนึงถึงบริบทของสถาบัน และสภาพการจัดการเรียนการสอน ดังนั้นข้อเสนอแนะควรมีความเป็นรูปธรรม เพราะข้อเสนอแนะที่เป็นนามธรรมมากเกินไป จะไม่สามารถปฏิบัติได้จริง ข้อเสนอแนะต้องมาจากผลการประเมินที่พัฒนามาจากบริบทของสถาบัน และสามารถเห็นผลจากการปฏิบัติได้

ประเด็นในการตรวจสอบรายงานผลการประเมินคุณภาพ (สุชาติดา บวรกิติวงศ์)

1. ในรายงานมีการอธิบายบริบทของสถาบัน ได้แก่ สภาพทั่วไป การดำเนินงาน ผลการดำเนินงาน เพียงพอหรือไม่

2. มีการอธิบายวิธีการดำเนินงานประเมิน การได้มาซึ่งการเก็บรวบรวมข้อมูลจากหลายแหล่งหรือไม่

3. ข้อมูลที่ได้มีความเพียงพอและมีความเป็นตัวแทนหรือไม่ มีความน่าเชื่อถือหรือไม่

4. รายงานผลการประเมินถูกต้องตามเกณฑ์ มีข้อค้นพบตรงไปตรงมาและมาจากข้อมูลที่รวบรวมหรือไม่
5. ข้อเสนอแนะมาจากผลการประเมิน เป็นรูปธรรมที่สถาบันสามารถนำไปปฏิบัติได้จริงหรือไม่
6. มีการระบุจุดแข็งพร้อมแนวทางเสริมให้ดีขึ้น ระบุจุดที่ควรพัฒนาพร้อมระบุวิธีการแก้ไขภายใต้ข้อมูลผลการประเมินและบริบทของสถาบันหรือไม่
7. การเขียนรายงานตรงตามรูปแบบที่กำหนด หรือไม่
8. เป็นรายงานที่มีการบูรณาการ ใช้ในการปฏิบัติงานจริง และมีประโยชน์หรือไม่
9. เป็นรายงานที่มีความเป็นปรนัย ที่ทุกคนอ่านสามารถเข้าใจได้ตรงกันหรือไม่
10. เป็นรายงานที่ได้รับการตรวจสอบทั้งจากทีมผู้ประเมินด้วยกันและจากผู้ประเมินอกิมาณภายในหรือไม่

6. บทบาทหน้าที่ของนักประเมินอกิมาณ

นักประเมินอกิมาณมีบทบาทหน้าที่ในการประเมินอกิมาณรายงานผลการประเมินคุณภาพภายใน ระดับสถาบันที่ผ่านการประเมินคุณภาพภายในจากผู้ประเมินคุณภาพภายใน ระดับอุดมศึกษา เพื่อตรวจสอบ ควบคุม และรับรองคุณภาพของผลการประเมิน เนื่องจากการประเมินอกิมาณจะให้สารสนเทศที่สำคัญ สองประการ คือ *ประการแรก* เป็นผลการวิเคราะห์วิธีวิทยาการประเมินที่ใช้อยู่ว่ามีความเหมาะสมมากน้อยเพียงใด และ*ประการที่สอง* เป็นผลการตัดสินคุณค่าว่าคุณภาพของสถาบันจากผลการประเมินเป็นอย่างไร และควรต้องมีการพัฒนาอย่างไร สำหรับใช้เป็นข้อมูลในการพิจารณาปรับปรุงกระบวนการประเมิน และการดำเนินงานของสถาบัน การประเมินอกิมาณครั้งนี้ใช้การประเมินอกิมาณแบบสรุปรวม (Summative evaluation) เนื่องจากการประเมินเมื่อสิ้นสุดการประเมินคุณภาพภายในแล้ว เพื่อแสดงให้เห็นให้ผู้มีส่วนได้ส่วนเสียและผู้ที่เกี่ยวข้องเห็นจุดแข็ง และจุดอ่อนของการประเมิน และตัดสินคุณค่าตามมาตรฐานการประเมินอกิมาณที่พัฒนาขึ้น

นักประเมินอกิมาณมีบทบาทในการประเมินอกิมาณ ดังรายละเอียดตามขั้นตอนการประเมินอกิมาณ ดังต่อไปนี้

ขั้นตอนในการประเมินอกิมาณ	บทบาทของนักประเมินอกิมาณ
1. การกำหนดจุดมุ่งหมายและประเภทของการประเมินอกิมาณ	เตรียมการประเมินอกิมาณแบบสรุปรวม (Summative evaluation) จากเอกสารรายงาน ได้แก่ SAR รายงานผลการประเมินคุณภาพภายใน และข้อมูลที่เกี่ยวข้องของสถาบันที่ทำการประเมินอกิมาณ

ขั้นตอนในการประเมินอภิमान	บทบาทของนักประเมินอภิमान
2. การกำหนดคณะนักประเมินอภิमान	ศึกษาข้อมูลพื้นฐาน SAR รายงานผลการประเมินคุณภาพภายใน ของสถาบันที่ได้รับมอบหมายร่วมกัน ภายในกลุ่มที่รับผิดชอบ หรือที่ประเมินสถาบันเดียวกัน
3. กำหนดมาตรฐาน เกณฑ์ การตัดสินคุณภาพการประเมิน	ศึกษามาตรฐาน เกณฑ์ในการตัดสินคุณภาพของรายงานผลการประเมินให้เข้าใจ เพื่อสามารถตัดสินผลการประเมินได้อย่างถูกต้อง
4. การสร้างเครื่องมือสำหรับการประเมินอภิमान	ศึกษาเครื่องมือที่ใช้ในการประเมินอภิमानอย่างเข้าใจ รวมถึงวิธีการประเมิน และเกณฑ์การให้คะแนน
5. การรวบรวมข้อมูล สารสนเทศที่มีอยู่	ดำเนินการรวบรวมข้อมูลจากเอกสาร หลักฐานที่เก็บรวบรวมมา และหากต้องการข้อมูลเพิ่มเติมให้ศึกษาจาก Website หรือ ติดต่อสถาบันที่ประเมิน
6. วิเคราะห์ข้อมูลสารสนเทศเชิงปริมาณ และเชิงคุณภาพ	วิเคราะห์สารสนเทศเชิงปริมาณ ด้วยค่าสถิติต่างๆ และวิเคราะห์สารสนเทศเชิงคุณภาพจากการเก็บรวบรวมข้อมูลทั้งหมด
7. ตัดสินงานประเมินตามมาตรฐาน เกณฑ์ การประเมิน	ดำเนินการประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน โดยนำผลการประเมินเทียบกับเกณฑ์ เพื่อตัดสินคุณค่าของการประเมิน
8. นำเสนอรายงานผลการประเมินอภิमान	นำเสนอผลการตัดสินประเมินอภิमान ฉบับร่าง เพื่อตรวจสอบความถูกต้อง และชัดเจนของการตัดสินผลการประเมิน
9.สรุปผลการประเมินและประยุกต์ใช้ผลการประเมินอภิमान	เขียนสรุปผลการประเมินอภิमान ให้ชัดเจน ถูกต้อง ครบถ้วน และนำเสนอผลการประเมินต่อผู้เกี่ยวข้อง

7. วิธีการประเมินอภิमान

วิธีการประเมินอภิमानของนักประเมินอภิमान ดำเนินการจากการสังเคราะห์ข้อค้นพบจากรายงานผลการประเมินคุณภาพภายใน รายงานการประเมินตนเอง หรือเอกสารอื่นๆ ของสถาบันประกอบ

ขั้นตอนของการประเมิน

ประกอบด้วย **ขั้นก่อนการประเมินอภิमान (1-4)** **ขั้นระหว่างการประเมินอภิमान (5-7)** และ**ขั้นหลังประเมินอภิमान (8-9)** ดังรายละเอียดต่อไปนี้

ขั้นก่อนการประเมินอภิमान (1-4) เป็นขั้นตอนในการวางแผนการประเมิน คือ การศึกษาข้อมูลพื้นฐาน ของสถาบัน และเตรียมการประเมินอภิमान จัดเตรียมคณะนักประเมิน เครื่องมือในการประเมิน ตามขั้นตอนต่อไปนี้

1. กำหนดจุดมุ่งหมายและประเภทของการประเมินอภิमान
2. กำหนดคณะนักประเมินอภิमान
3. กำหนดมาตรฐาน เกณฑ์ การตัดสินคุณภาพการประเมิน
4. การสร้างเครื่องมือ สำหรับการประเมินอภิमान

ขั้นระหว่างการประเมินอภิमान (5-7) เป็นขั้นตอนในการศึกษารายงานผลการประเมิน และรวบรวมข้อมูลของสถาบันในเชิงลึก เพื่อเข้าใจ และสามารถตัดสินงานประเมินตามมาตรฐาน เกณฑ์การประเมิน

5. การรวบรวมข้อมูล สารสนเทศที่มีอยู่
6. วิเคราะห์ข้อมูลสารสนเทศเชิงปริมาณและเชิงคุณภาพ
7. ตัดสินงานประเมินตามมาตรฐาน เกณฑ์การประเมิน

ขั้นหลังประเมินอภิमान (8-9) เป็นขั้นตอนจัดทำรายงานผลการประเมินอภิमान ตามข้อมูลที่รวบรวม และตัดสินผลการประเมิน เพื่อเสนอแนะให้สถาบันนำไปใช้ประโยชน์ต่อไป

8. นำเสนอรายงานผลการประเมินอภิमान
9. สรุปผลการประเมินและประยุกต์ใช้ผลการประเมินอภิमान

8. คุณสมบัติของนักประเมินอภิमान (สุวิมล ว่องวาณิช, 2550)

1. เป็นผู้มีความซื่อสัตย์ ยึดมั่นในจรรยาบรรณของการประเมิน ทำงานเพื่อประโยชน์ส่วนร่วม รักความยุติธรรม ทำงานด้วยความโปร่งใส ตรวจสอบได้
2. มีความรู้ ความสามารถ ทักษะ ประสบการณ์ด้านการประเมินและความเชื่อในคุณค่าของการประเมิน
3. มีความสามารถในการทำงานแบบมีส่วนร่วมและการมีคุณสมบัติส่วนบุคคลที่ทำให้สามารถเข้าถึงข้อมูลที่นักประเมิน (ปฐมภูมิ) ใช้ในการประเมิน
4. มีความรู้ความเข้าใจในบริบท เงื่อนไขจำกัดของสิ่งที่ถูกประเมินและการปฏิบัติงานด้านการประเมินของนักประเมินปฐมภูมิ
5. ความสามารถในการสืบค้น วิเคราะห์ วิวิจารณ์ จุดแข็งจุดอ่อน ของการออกแบบการประเมิน กระบวนการประเมิน และสรุปผลการประเมิน

6. ความสามารถในการสื่อสารผลการประเมินอภิमानในทางสร้างสรรค์ การให้ข้อเสนอแนะที่นำไปสู่การพัฒนาปรับปรุงงานการประเมินในอนาคต

9. ส่วนประกอบของรายงานการประเมินอภิमान

รายงานการประเมินอภิमानประกอบด้วย 3 ส่วน คือ ส่วนนำ ส่วนสาระการประเมิน และ ส่วนสรุปและเสนอแนะ มีรายละเอียดดังต่อไปนี้

ส่วนที่ 1 ส่วนนำ

ประกอบด้วย

- ปก
- คำนำ
- สารบัญ
- บทสรุปผลการประเมินอภิमान ได้แก่ ผลการประเมินในภาพรวมของแต่ละมาตรฐาน จุดเด่น จุดด้อย

ส่วนที่ 2 ส่วนสาระการประเมิน

ประกอบด้วย

- วิธีการประเมิน
- ผลการประเมินทั้งภาพรวม และรายมิติ
- ผลการประเมินรายมาตรฐาน โดยนำเสนอเป็นตารางตามแบบประเมินอภิमान

ส่วนที่ 3 ส่วนสรุปและเสนอแนะ

ประกอบด้วย

- ความสำคัญโดยรวม
- จุดแข็ง จุดอ่อนของการประเมินภายใน
- ข้อเสนอแนะในการนำผลการประเมินอภิमानไปใช้ประโยชน์
- เหตุผลสนับสนุนการตัดสินใจ

แบบประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน

ตอนที่ 1 ข้อมูลทั่วไปของสถาบันอุดมศึกษา

คำชี้แจง โปรดทำเครื่องหมาย ✓ ในช่อง ตามข้อมูลจริง

- ชื่อมหาวิทยาลัย/วิทยาลัย.....
- ประเภทสถาบันอุดมศึกษา
 สถาบันอุดมศึกษารัฐบาล/ในกำกับ สถาบันอุดมศึกษาเอกชน มหาวิทยาลัยราชภัฏ มหาวิทยาลัยราชภัฏมณฑล
- ที่ตั้งสถาบันอุดมศึกษา
 กรุงเทพมหานคร ภาคเหนือ ภาคกลาง ภาคตะวันออกเฉียงเหนือ ภาคใต้

ตอนที่ 2 แบบประเมินอภิमानรายงานผลการประเมินคุณภาพภายใน

เกณฑ์การแปลความหมายผลการประเมิน ตามระดับคุณภาพ ดังนี้

- | | |
|---------------|-------------------------------------|
| 3.26 ขึ้นไป | หมายถึง การประเมินมีคุณภาพดีมาก |
| 2.51 – 3.25 | หมายถึง การประเมินมีคุณภาพดี |
| 1.76 – 2.50 | หมายถึง การประเมินมีคุณภาพพอใช้ |
| น้อยกว่า 1.76 | หมายถึง ควรปรับปรุงคุณภาพการประเมิน |

ตารางสรุปคะแนนประเมินอภิमानรายมาตรฐาน

มาตรฐานที่	มาตรฐานการประเมิน	น้ำหนัก (A)	ผลรวมของผลการประเมิน $\Sigma (AXB)$	คะแนนเฉลี่ยผลการประเมิน $(\Sigma (AXB)/A)$	แปลความหมาย
1.	ความตรงของการประเมิน (Validity)	30			
2.	อรรถประโยชน์ (Utility)	30			
3.	จริยธรรมในการประเมิน (Ethicality)	24			
4.	ความน่าเชื่อถือของนักประเมิน (Credibility)	8			
5.	ประสิทธิผลต้นทุน (Cost-effectiveness)	8			
	ค่าเฉลี่ยถ่วงน้ำหนัก	100			

คำชี้แจง ให้พิจารณาข้อมูลสารสนเทศที่เก็บรวบรวมมาว่าสอดคล้องกับระดับเกณฑ์การประเมินอภิमानในระดับใด โดยใส่หมายเลข 4 3 2 1 ลงในช่องคะแนนการประเมิน

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
ความตรงของการประเมิน (Validity) 13 ตัวบ่งชี้	30							
Va 1.1 การระบุบริบทของสถาบันที่ส่งผลกระทบต่อ การประเมิน	2	มีการระบุประวัติ ข้อมูลพื้นฐาน สภาพทั่วไป การดำเนินงาน และ ผลการดำเนินงาน ของสถาบันที่ ประเมินอย่างชัดเจน รวมถึง วิเคราะห์ผลกระทบต่อ การ ดำเนินการ และผลผลิตของ สถาบัน	มีการระบุประวัติ ข้อมูลพื้นฐาน และมีการอธิบายถึงสภาพทั่วไป การดำเนินงาน และ ผลการ ดำเนินงาน ของสถาบันที่ประเมิน อย่างชัดเจน	มีการระบุประวัติและข้อมูล พื้นฐาน ของสถาบันใน รายงานการประเมิน	ไม่ระบุประวัติหรือข้อมูล พื้นฐาน ของสถาบันใน รายงานการประเมิน			
Va 1.2 การระบุถึงอัตลักษณ์ ของสถาบัน	2	มีการระบุถึงอัตลักษณ์หรือจุดเน้น ของสถาบัน และมีกิจกรรม ดำเนินการตามแผน เพื่อให้ สามารถบรรลุจุดเน้นดังกล่าว	มีการระบุถึงอัตลักษณ์หรือจุดเน้น ของสถาบันและมีการกำหนด นโยบาย แผนการประเมินตาม จุดเน้นดังกล่าว	มีการระบุถึงอัตลักษณ์หรือ จุดเน้นของสถาบันไว้ใน รายงานการประเมิน	ไม่มี การระบุถึงอัตลักษณ์ หรือจุดเน้นของสถาบันไว้ใน รายงานการประเมิน			
Va 1.3 การบรรยายถึง วัตถุประสงค์ในการประเมิน อย่างชัดเจน	2	มีการระบุถึงวัตถุประสงค์เฉพาะไว้ใน รายงาน อย่างชัดเจน สามารถ ตรวจสอบได้ และอธิบายแนวทางการ ประเมิน เพื่อให้บรรลุ วัตถุประสงค์	มีการระบุวัตถุประสงค์ในการ ประเมินอย่างชัดเจน ไว้ในรายงาน การประเมิน	มีการระบุวัตถุประสงค์ทั่วไป ในการประเมินไว้ในรายงาน การประเมิน	ไม่มี การระบุวัตถุประสงค์ ในการ ประเมินไว้ใน รายงาน			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Va 1.4 การออกแบบการประเมินที่สอดคล้องกับสภาพ และบริบทของสถาบัน	3	มีการวางแผนและออกแบบการประเมินได้แก่ แหล่งข้อมูล การเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล การแปลความหมายและเขียนรายงาน ที่สอดคล้องกับสภาพ บริบทของสถาบัน	มีการวางแผน ออกแบบ การประเมินได้แก่ แหล่งข้อมูล การเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล การแปลความหมายและเขียนรายงานไว้ในขั้นตอนการเตรียมการและวางแผนก่อนการประเมิน	มีการวางแผนและออกแบบการประเมินเฉพาะ การเก็บรวบรวมข้อมูล หรือการวิเคราะห์ข้อมูลไว้ ก่อนการประเมิน	ไม่มีรายละเอียดขั้นตอนเตรียมการวางแผนการประเมิน			
Va 1.5 การวิเคราะห์แหล่งความแตกต่างของเอกสารในการประเมิน	2	มีรายละเอียดการรวบรวมเอกสารการประเมินอย่างถูกต้องและชัดเจน โดยแสดงผลการวิเคราะห์ความแตกต่างของเอกสารที่มาจากแหล่งต่างกัน และรายงานผลการประเมินอย่างถูกต้องเหมาะสม	มีรายละเอียดการรวบรวมเอกสารการประเมินชัดเจน ระบุที่มาของเอกสารจากหลายแหล่ง แต่ไม่มีผลการวิเคราะห์ความแตกต่างของเอกสาร	มีรายละเอียดการรวบรวมเอกสารประกอบในการประเมิน และรายงานผลการประเมิน	เอกสารในการประเมิน ใช้เพียงแหล่งเดียว คือ รายงานการประเมินตนเอง (SAR)			
Va 1.6 การกำหนดแหล่งข้อมูลที่มีความน่าเชื่อถือ	2	มีการระบุการเก็บรวบรวมข้อมูลหลายแหล่ง โดย ตรวจสอบแบบ สุ่ม เปรียบเทียบ แหล่งข้อมูลแต่ละแหล่งที่เก็บรวบรวมมาสอดคล้องกับเกณฑ์มาตรฐานการประเมิน	มีกระบวนการเก็บรวบรวมข้อมูลหลายแหล่งทั้งการตรวจเยี่ยม สัมภาษณ์ และศึกษาเอกสาร	ใช้แหล่งข้อมูลเพียง 2 แหล่ง ได้แก่ การศึกษาเอกสาร การสัมภาษณ์ ในการเก็บรวบรวมข้อมูล และมีการรายงานไว้	ใช้แหล่งข้อมูลจาก การศึกษาเพียงแหล่งเดียว ในการเก็บรวบรวมข้อมูล			
Va 1.7 การตรวจสอบความถูกต้องของข้อมูล	3	มีการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา และมีหลักฐานยืนยันความสอดคล้องกันของข้อมูลต่างๆ มีรายละเอียดชัดเจน	มีการระบุการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา แต่หลักฐานยืนยันความสอดคล้องกันของข้อมูลต่างๆ ยังไม่มีรายละเอียด	มีการระบุการตรวจสอบความถูกต้องของข้อมูลที่เก็บรวบรวมมา แต่ไม่มีหลักฐานยืนยันแสดงถึงความสอดคล้องกันของข้อมูล	ไม่ระบุการตรวจสอบความถูกต้องของข้อมูล			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Va 1.8 การตรวจสอบคุณภาพของสารสนเทศ และตรงกับความต้องการของผู้ใช้ข้อมูล	2	มีการตรวจสอบคุณภาพความถูกต้อง สมบูรณ์ของสารสนเทศ สอดคล้องและตรงกับความต้องการของผู้ใช้ข้อมูลส่วนมาก	มีการระบุงการตรวจสอบสารสนเทศที่เก็บรวบรวมมากับความต้องการของผู้ใช้ข้อมูล	มีการตรวจสอบคุณภาพการเก็บรวบรวมข้อมูลสารสนเทศบางส่วน	ไม่มี การตรวจสอบคุณภาพของสารสนเทศที่เก็บรวบรวมมาได้			
Va 1.9 การวิเคราะห์ข้อมูลและประมวลผลข้อมูลอย่างถูกต้องตามวิธีการที่กำหนด	3	มีการวิเคราะห์ข้อมูลทั้งเชิงปริมาณและเชิงคุณภาพอย่างเป็นระบบ มีความถูกต้อง น่าเชื่อถือ ทุกตัวบ่งชี้ สามารถประมวลผลข้อมูลได้อย่างมีประสิทธิภาพ	การวิเคราะห์ข้อมูลเชิงปริมาณและเชิงคุณภาพ มีความถูกต้องอย่างน้อย 39 ตัวบ่งชี้ และมีหลักฐานยืนยันความสอดคล้องของข้อมูลต่างๆ	การวิเคราะห์ข้อมูลเชิงปริมาณและเชิงคุณภาพ มีความถูกต้อง 36 ตัวบ่งชี้ ขึ้นไป	มีการวิเคราะห์ข้อมูลเฉพาะข้อมูลเชิงปริมาณเท่านั้น แต่ยังมี ความถูกต้อง น้อยกว่า 36 ตัวบ่งชี้			
Va 1.10 การแปลผลและสรุปผลการประเมินถูกต้องตามหลักฐานที่ปรากฏ	3	การแปลผลและสรุปผลการประเมินอย่างถูกต้อง ชัดเจน มีการอ้างอิงสารสนเทศสนับสนุนผลการประเมินทุกครั้ง และสอดคล้องกับผลการวิเคราะห์ข้อมูล	การแปลผลและสรุปผลการประเมินถูกต้อง สอดคล้องกับผลการวิเคราะห์ข้อมูล และมีการอ้างอิงสารสนเทศที่สนับสนุนผลการประเมินบางส่วน	การแปลผลและสรุปผลการประเมิน สอดคล้องกับผลการวิเคราะห์ข้อมูล แต่ยังไม่มีการอ้างอิงสารสนเทศที่สนับสนุนผลการประเมิน	การแปลผลและสรุปผล ไม่สอดคล้องกับผลการประเมิน และยังขาดการอ้างอิงสารสนเทศที่สนับสนุนผลการประเมิน			
Va 1.11 การเขียนรายงานผลตามข้อค้นพบทั้งเชิงบวกและเชิงลบ	2	การเขียนรายงานผลการประเมินอย่างถูกต้อง ทั้งเชิงบวกและเชิงลบ โดยเสนอทั้งจุดเด่นและแนวทางการพัฒนาการประเมินชัดเจน	การเขียนรายงานผลการประเมินเน้นจุดเด่น แต่มีการนำเสนอทั้งจุดเด่น และ แนวทางการพัฒนาการประเมินบางส่วน	การเขียนรายงานผลการประเมิน เสนอเฉพาะข้อค้นพบด้านเดียว เป็นเชิงบวกหรือเชิงลบเท่านั้น	การเขียนรายงานผลการประเมิน ไม่มี การนำเสนอจุดเด่นหรือจุดที่ควรพัฒนา ของการประเมิน			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Va 1.12 การรายงานผลการประเมินอย่างเป็นกลาง	2	มีการรายงานผลการประเมินอย่างชัดเจน ตรวจสอบได้ โดยผลการประเมินสามารถรองรับการตัดสินใจของทุกกลุ่มที่เกี่ยวข้อง	มีการรายงานผลการประเมินอย่างเป็นกลาง ไม่มี หลักฐานแสดงว่าผลการประเมินถูกบิดเบือน และระบุว่าผู้ที่เกี่ยวข้องกับการประเมิน ส่วนใหญ่ร่วมพิจารณาตัดสินใจ	มีการระบุว่า รายงานผลการประเมินอย่างเป็นกลาง แต่ ไม่มี หลักฐานว่าพิจารณาจากการตัดสินใจของกลุ่มผู้ที่เกี่ยวข้องกับการประเมิน	ไม่มี หลักฐานแสดงว่า การรายงานผลการประเมินอย่างเป็นกลาง			
Va 1.13 การตรวจสอบความถูกต้องของรายงานการประเมินจากผู้ที่เกี่ยวข้อง	2	มีการตรวจสอบความถูกต้อง นำเชื่อถือของผลการประเมิน และประเมินจุดแข็ง จุดอ่อนของการประเมิน โดยคณะผู้ประเมินเอง หรือผู้ที่มีส่วนเกี่ยวข้องกับการประเมิน	มีการตรวจสอบความถูกต้อง ของผลการประเมินตามมาตรฐานการประเมินที่กำหนดขึ้น	มีการกำหนดแนวทางเพื่อตรวจสอบผลการประเมิน โดยผู้ที่มีส่วนเกี่ยวข้องกับการประเมิน	ไม่มี การรายงานการตรวจสอบผลการประเมินจากผู้ที่เกี่ยวข้องกับการประเมิน			
อรรถประโยชน์ (Utility) 10 ตัวบ่งชี้	30							
Ut 2.1 การระบุผู้มีส่วนได้ส่วนเสียที่เกี่ยวข้องในการประเมิน	3	มีการระบุถึงผู้มีส่วนได้ส่วนเสียครบทุกกลุ่ม เข้ามามีส่วนร่วมและเกี่ยวข้องกับการประเมิน	มีการระบุผู้มีส่วนได้ส่วนเสียอย่างน้อย 2 กลุ่ม เข้ามามีส่วนร่วมในการประเมิน	มีการระบุผู้มีส่วนได้ส่วนเสียที่มีส่วนร่วมในการประเมินเพียงกลุ่มเดียว	ไม่มี การระบุผู้มีส่วนได้ส่วนเสียในการประเมิน			
Ut 2.2 การระบุช่วงเวลาและระยะเวลาในการเก็บรวบรวมข้อมูลที่เหมาะสม	3	มีการระบุ ช่วงเวลา ระยะเวลา และรายละเอียดในการเก็บรวบรวมข้อมูล อย่างชัดเจน เพื่อให้ผู้รับผลการประเมินสามารถจัดเตรียมข้อมูล ให้สอดคล้องกับการประเมิน	มีการระบุ ช่วงเวลาเพียงกว้างๆ ในการเก็บรวบรวมข้อมูล แต่ไม่ระบุรายละเอียดของระยะเวลาในการเก็บรวบรวมข้อมูลได้	มีการระบุกำหนดการในการเก็บรวบรวมข้อมูลไว้ เป็นประมาณการระยะเวลา ในการเก็บรวบรวมข้อมูลเท่านั้น	ไม่มี การระบุ แผนเกี่ยวกับช่วงเวลา และระยะเวลา ในการเก็บรวบรวมข้อมูล			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Ut 2.3 การกำหนดเทคนิคในการเก็บรวบรวมข้อมูล โดยคำนึงถึงข้อจำกัดของทรัพยากรการประเมิน	3	มีการวางแผนกำหนดเทคนิคการประเมินตามทรัพยากรที่มีอย่างจำกัด โดยคำนึงถึงผลกระทบต่อผลการประเมิน และข้อจำกัดในการประเมิน	มีการวางแผนกำหนดเทคนิคการประเมินไว้ ตามทรัพยากรที่มีอย่างจำกัด	มีการวางแผนกำหนดเทคนิคในการประเมินไว้ล่วงหน้า	ไม่มีมีการวางแผนเกี่ยวกับการกำหนดเทคนิคในการประเมินไว้ล่วงหน้า			
Ut 2.4 การตรวจสอบการตัดสินใจผลการประเมิน สอดคล้องกับสภาพจริง	3	มีการพิจารณาการตัดสินใจผลการประเมินที่ชัดเจนตามมาตรฐานการประเมิน ผลการวิเคราะห์โดยอ้างอิงหลักฐานที่น่าเชื่อถือตามสภาพจริง	มีการพิจารณาการตัดสินใจผลการประเมินตามมาตรฐานการประเมิน ผลการวิเคราะห์และพยายามอ้างอิงหลักฐานประกอบ แต่ยังไม่ชัดเจน	มีการพิจารณาตัดสินใจผลการประเมิน ตามมาตรฐานการประเมิน แต่ไม่มีอ้างอิงหลักฐาน	ไม่มี การระบุพิจารณาการตัดสินใจผลการประเมินตามมาตรฐานการประเมิน			
Ut 2.5 ผลการประเมินเป็นประโยชน์ ชี้ให้เห็นทิศทางการปรับปรุง การพัฒนาการดำเนินงานของสถาบัน	3	ผลการประเมินเป็นประโยชน์ในการวางแผน การดำเนินงาน โดยข้อเสนอแนะจากการประเมิน ชี้ให้เห็นทิศทางการพัฒนา มีการประชุมเพื่อรับฟังข้อมูลย้อนกลับจากการใช้ผลการประเมิน	ผลการประเมินเป็นประโยชน์ในการวางแผนการดำเนินงาน ปรับปรุงตามข้อเสนอแนะ แต่ไม่มีการประชุมเพื่อรับฟังข้อมูลย้อนกลับจากการใช้ผลการประเมิน	ผลการประเมินเป็นประโยชน์ในการวางแผนการดำเนินงาน การดำเนินงาน แต่ข้อเสนอแนะยังไม่สามารถนำไปปฏิบัติได้	ผลการประเมิน ไม่มีการระบุ ข้อเสนอแนะที่นำไปใช้ประโยชน์ และปรับปรุงการดำเนินงาน			
Ut 2.6 การเขียนรายงาน สอดคล้องกับรูปแบบที่กำหนด	3	การเขียนรายงานการประเมิน ถูกต้อง ครบถ้วน ทั้งรายชื่อคณะผู้ประเมิน บทนำ วิธีการประเมิน ผลการประเมินและข้อเสนอแนะ ภาคผนวก ตามรูปแบบที่สกอ. กำหนด จัดพิมพ์ถูกต้อง	การเขียนรายงานการประเมิน ครบถ้วน ทั้งรายชื่อคณะผู้ประเมิน บทนำ วิธีการประเมิน ผลการประเมินและ ข้อเสนอแนะ ภาคผนวก ตามรูปแบบที่สกอ.กำหนด อาจมีการพิมพ์และจัดหน้าผิดพลาดเล็กน้อย	การเขียนรายงานการประเมิน เป็นไปตามรูปแบบที่กำหนด แต่ขาดยังส่วนสาระสำคัญ บางส่วน ได้แก่ ส่วนสรุป หรือวิธีการประเมิน หรือ ข้อเสนอแนะ	การเขียนรายงานยังไม่ ครบถ้วน และไม่ถูกต้องตามรูปแบบที่สกอ.กำหนดไว้			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Ut 2.7 รายงานผลการประเมินครบถ้วน ครอบคลุมผลการประเมิน และข้อเสนอแนะเพื่อการพัฒนา	3	รายงานการประเมินมีเนื้อหาครบถ้วนตามที่กำหนด ครอบคลุมผลการประเมิน เสนอจุดเด่น จุดที่ควรพัฒนา ระบุสาเหตุของปัญหา และข้อเสนอแนะที่เป็นประโยชน์นำไปปฏิบัติได้	รายงานการประเมินมีเนื้อหาครบถ้วนตามที่กำหนด ครอบคลุมผลการประเมิน มีการเสนอจุดเด่น จุดที่ควรพัฒนา แต่ไม่ระบุสาเหตุของปัญหา และข้อเสนอแนะนำไปปฏิบัติได้ยาก	รายงานการประเมินมีเนื้อหาครบถ้วน ครอบคลุมผลการประเมิน จุดเด่น จุดที่ควรพัฒนา บางองค์ประกอบ แต่ขาดข้อเสนอแนะที่เป็นประโยชน์	รายงานการประเมินยังมีเนื้อหาไม่ครบถ้วน ครอบคลุมผลการประเมิน			
Ut 2.8 รายงานผลการประเมินเขียนด้วยภาษาเชิงวิชาการ และมีสื่อประกอบ	3	การเขียนรายงานการประเมิน ใช้ภาษาเชิงวิชาการที่ชัดเจน เข้าใจง่าย และมีการนำเสนอด้วยสื่อ เช่น รูปภาพ กราฟ ที่น่าสนใจประกอบ	การเขียนรายงานการประเมิน ใช้ภาษาที่ชัดเจน เน้นข้อความตัวอักษร มีสื่อที่เป็นตารางประกอบบางส่วน	การเขียนรายงานการประเมิน ใช้ภาษาที่ค่อนข้างชัดเจน แต่ไม่มีสื่อ ประเภทตาราง ข้อมูลประกอบ	การเขียนรายงานการประเมิน ใช้ภาษาที่ไม่ชัดเจน อ่านไม่เข้าใจ และไม่มีสื่อประกอบ			
Ut 2.9 การจัดทำรายงานผลการประเมินเสร็จทันเวลา	3	การจัดทำรายงานการประเมิน มีข้อมูลที่สมบูรณ์ เสร็จทันเวลาที่กำหนด ตรงช่วงเวลาที่ผู้ผลการประเมินต้องการ	มีการจัดทำรายงานการประเมินเสร็จทันเวลาที่กำหนด แต่ยังขาดความสมบูรณ์ในบางส่วน	การจัดทำรายงานการประเมินมีข้อมูลที่สมบูรณ์ แต่ยังเสร็จล่าช้ากว่ากำหนด (30 ก.ย.)	การส่งรายงานการประเมินล่าช้า กว่ากำหนด (30 ก.ย.)			
Ut 2.10 การเผยแพร่ผลการประเมินสู่ผู้ใช้ผลการประเมินทุกกลุ่ม	3	มีการเผยแพร่ผลการประเมินอย่างเปิดเผย ด้วยภาษาที่เหมาะสมสื่อความได้ต่อความเข้าใจกับทุกคน	มีการเผยแพร่ผลการประเมิน ด้วยภาษาเขียนที่สามารถเข้าใจได้ง่าย	มีการเผยแพร่ผลการประเมิน ด้วยภาษาที่กำกวม สามารถเข้าใจได้เฉพาะกลุ่ม	มีการเผยแพร่ผลการประเมินใช้ภาษาพูด เน้นเฉพาะกลุ่มผู้รับการประเมิน			
จริยธรรมในการประเมิน (Ethicality) 9 ตัวบ่งชี้	24							

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Et 3.1 การสื่อสารกับผู้มีส่วนร่วมในการประเมินตลอดช่วงเวลาประเมิน	2	มีการสื่อสารกำหนดวิธีการประเมิน คณะผู้ประเมินที่เหมาะสมกับสภาพความเป็นจริง โดยมีการเจรจากับผู้มีส่วนร่วมในการประเมิน ตลอดช่วงระยะเวลาทำการประเมิน	มีการสื่อสารกำหนดวิธีการประเมิน คณะผู้ประเมิน ที่สอดคล้องกับสภาพความเป็นจริง แต่ ไม่มี หลักฐานที่ชัดเจนในการเจรจากับผู้มีส่วนร่วมในการประเมิน	มีการสื่อสารวางแผนกำหนดวิธีการประเมิน แต่ยังไม่สอดคล้องกับสภาพจริง ทำให้เป็นภาระของ สถาบันในการประเมิน	ไม่มี การสื่อสารการวางแผนกำหนดวิธีการประเมินตามสภาพความเป็นจริง			
Et 3.2 ผลการประเมินเป็นที่ยอมรับจากผู้ที่เกี่ยวข้อง และปราศจากอคติทางการประเมิน	2	ผลการประเมินเป็นที่ยอมรับจากบุคคลหลายกลุ่ม ปราศจากอคติที่ขัดขวางการประเมิน ซึ่งผลการประเมินสะท้อนให้เห็นจุดเด่น จุดที่ควรพัฒนา และทิศทางการพัฒนาในอนาคต	การประเมิน เป็นไปตามวัตถุประสงค์และเกณฑ์การประเมิน มีความพยายามประสานงานเพื่อเชื่อมโยงความเข้าใจกับบุคคลหลายกลุ่ม แต่มีการหักท้วงผลการประเมินด้วยวาจา	การประเมิน บรรลุวัตถุประสงค์ เป็นไปตามเกณฑ์การประเมิน แต่มีการหักท้วงผลการประเมินเป็นลายลักษณ์อักษรจากผู้รับการประเมิน	ผลการประเมิน ไม่ได้รับ การยอมรับจากบุคคลหลายกลุ่ม			
Et 3.3 การปรับปรุงคุณภาพการประเมินอย่างต่อเนื่อง โดยคำนึงถึงความต้องการจำเป็นของผู้รับการประเมิน	3	มีการพิจารณาปรับปรุงคุณภาพผลการประเมิน ระหว่างการประเมิน และอธิบายเปรียบเทียบผลการประเมินกับความต้องการของผู้รับการประเมิน	มีการพิจารณาปรับปรุงคุณภาพผลการประเมิน ระหว่างการประเมิน แต่ยังไม่ครอบคลุมความต้องการจำเป็นของผู้รับการประเมิน	มีการทบทวนวางแผนการประเมิน และพิจารณาปรับปรุงคุณภาพผลการประเมิน แต่มีรายละเอียดไม่ชัดเจน	ไม่มี การทบทวนแผนการประเมิน และพิจารณาปรับปรุงคุณภาพผลการประเมิน			
Et 3.4 การจัดทำเอกสารข้อตกลงการประเมิน	2	มีการดำเนินการประเมิน ตามกำหนดการและวิธีการประเมิน ที่สอดคล้องกับวัตถุประสงค์ กระบวนการ และทรัพยากรในการประเมิน	มีการตรวจสอบ ทบทวน กำหนดการและวิธีการประเมิน ที่สอดคล้องกับทรัพยากรในการประเมิน	มีการระบุกำหนดการประเมิน และวิธีการประเมินระหว่างผู้ประเมินกับสถาบัน	มีการจัดทำกำหนดการประเมินระหว่างผู้ประเมินกับสถาบัน			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Et 3.5 การเปิดเผยผลการประเมินและข้อจำกัดที่มีผลกระทบต่อการประเมิน	3	ผู้ที่เกี่ยวข้องรับทราบผลการประเมินและข้อจำกัดที่ส่งผลกระทบต่อการประเมิน อย่างเปิดเผย ชัดเจน ครบถ้วนทุกองค์ประกอบ	ผู้ที่เกี่ยวข้องรับทราบผลการประเมินและข้อจำกัดที่ส่งผลกระทบต่อการประเมิน แต่ยังมีรายละเอียดไม่ชัดเจน ครอบคลุมทุกองค์ประกอบ	ผู้ที่เกี่ยวข้องรับทราบเฉพาะผลการประเมิน แต่ ไม่มี การเปิดเผยข้อจำกัดที่ส่งผลกระทบต่อการประเมิน	ไม่มี การเปิดเผยผลการประเมินและข้อจำกัดที่ส่งผลกระทบต่อการประเมินให้ผู้ที่เกี่ยวข้องรับทราบ			
Et 3.6 การกำหนดวิธีการประเมิน โดยคำนึงถึงความปลอดภัย รักษาความลับ และ ไม่ก่อให้เกิดความเสียหายกับผู้ให้ข้อมูล	3	ไม่มี การเปิดเผย ข้อมูลความลับของผู้มีส่วนร่วมในการประเมิน กรณีที่สร้างความเสียหายหรืออันตรายโดยตรงได้	มีวิธีการรักษาข้อมูลความลับ เช่น ไม่ระบุคำพูด ของผู้มีส่วนร่วมในการประเมิน กรณีที่สามารถสร้างความเสียหายหรืออันตรายได้	มีวิธีการรักษาข้อมูลความลับ โดยไม่ระบุตัวตน ของผู้มีส่วนร่วมในการประเมิน กรณีที่อาจสร้างความเสียหายหรืออันตรายได้	มีการเปิดเผย ข้อมูลความลับ ที่ก่อให้เกิดความเสียหาย และอันตรายแก่ผู้มีส่วนร่วมในการประเมิน			
Et 3.7 การติดต่อกับผู้ที่เกี่ยวข้อง โดยคำนึงถึงความแตกต่าง หลากหลายของผู้ที่เกี่ยวข้อง	3	มีการตรวจสอบภูมิหลังทางของผู้ที่เกี่ยวข้องกับการประเมิน และ ดำเนินการประเมินโดยมีการสื่อสาร ข้อมูล และ สร้างความสัมพันธ์กับผู้ที่เกี่ยวข้องในการประเมินพื้นฐานแตกต่างกัน	มีการศึกษาภูมิหลังของผู้ที่เกี่ยวข้องกับการประเมิน และ พยายามสร้างความสัมพันธ์กับผู้ที่เกี่ยวข้องกับการประเมินที่มีพื้นฐานแตกต่างกัน	มีการสอบถามถึงภูมิหลังของผู้ที่เกี่ยวข้องกับการประเมิน แต่ยัง ไม่มี การสื่อสารและสร้างความสัมพันธ์กับผู้ที่เกี่ยวข้องกับการประเมิน	ไม่มี ข้อมูลภูมิหลังของผู้ที่เกี่ยวข้องกับการประเมิน ได้แก่ นักศึกษา ศิษย์เก่า ผู้ปกครอง อาจารย์ และ ผู้ใช้บัณฑิต เป็นต้น			
Et 3.8 การประเมินที่เป็นธรรมต่อผู้ที่เกี่ยวข้องทุกฝ่าย	3	มีการตัดสินผลการประเมินจากฐานข้อมูลที่ต้อง ครบถ้วน ตรวจสอบจุดเด่น จุดที่ควรพัฒนา โดยให้ข้อเสนอแนะที่เป็นกลางกับผู้ให้นำผลการประเมินไปใช้	มีการตรวจสอบความถูกต้องของฐานข้อมูลก่อนการตัดสินผลการประเมิน มีความพยายามพิจารณาข้อเสนอแนะที่เป็นกลางต่อผู้ใช้งานการประเมิน	มีการตรวจสอบความถูกต้องของฐานข้อมูลก่อนการตัดสินผลการประเมิน แต่ ไม่ได้ พิจารณาข้อเสนอแนะที่เป็นกลางต่อผู้ใช้งานการประเมิน	ไม่มี การตรวจสอบความถูกต้องของการตัดสินผลการประเมิน			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Et 3.9 การประเมินที่ครบถ้วน ชัดเจนสอดคล้องกับมาตรฐาน และ ตัวบ่งชี้	3	มีการเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล และสรุปข้อมูลรายงานผลการประเมินถูกต้อง ครบถ้วน สอดคล้องกันทุกองค์ประกอบและมีข้อเสนอแนะที่เป็นประโยชน์และสอดคล้องกันตลอด	มีการเก็บรวบรวมข้อมูล วิเคราะห์ข้อมูล และสรุปข้อมูลรายงานผลการประเมินถูกต้อง ครบถ้วน สอดคล้องกับองค์ประกอบแต่ ข้อเสนอแนะยังไม่ชัดเจน นำไปปฏิบัติได้ยาก	มีการเก็บรวบรวมข้อมูล การวิเคราะห์ ข้อมูล สอดคล้องกับการสรุปข้อมูล นำไปใช้ประโยชน์ ในบาง องค์ประกอบ	ไม่มีการ ตรวจสอบความถูกต้องของข้อมูล การวิเคราะห์ข้อมูล และความ สอดคล้องของการสรุป ข้อมูล นำไปใช้ประโยชน์			
ความน่าเชื่อถือของนักประเมิน (Credibility) 4 ตัวบ่งชี้	8							
Cr 4.1 ผู้ประเมินมีความสามารถในการประเมิน	2	ผู้ประเมินวิเคราะห์ผลการประเมิน และเสนอรายงานผลการประเมินอย่างโปร่งใส ละเอียด ชัดเจน สอดคล้องกับสรุปผล และมีฐานข้อมูลรองรับ ส่วนใหญ่	ผู้ประเมินวิเคราะห์ผลการประเมิน และเสนอรายงานผลการประเมินอย่างชัดเจน สอดคล้องกับสรุปผล และมีฐานข้อมูลรองรับ บางส่วน	ผู้ประเมินวิเคราะห์ผลการประเมินและเสนอรายงานผลการประเมินไม่ชัดเจน การสรุปผลไม่มีฐานข้อมูลรองรับ	ผู้ประเมินวิเคราะห์ ผล การประเมิน และเขียน รายงานขัดแย้งกัน ตลอด เล่ม			
Cr 4.2 ผู้ประเมินมีทักษะในการสื่อสารผลการประเมิน	2	ผู้ประเมินสามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน ที่ชัดเจนทำให้เห็นทิศทางในการวางแผน และพัฒนาอย่างเหมาะสม นำไปปฏิบัติได้	ผู้ประเมินสามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน ครอบคลุม ผลการประเมิน แต่ยังมีรายละเอียดกว้างๆ นำไปปฏิบัติได้ บางส่วน	ผู้ประเมินสามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน แต่ไม่สามารถปฏิบัติเป็นรูปธรรมได้	ผู้ประเมิน ไม่ สามารถเสนอแนวทางเสริมจุดแข็ง และปรับปรุงจุดอ่อน ที่เป็นประโยชน์ได้			
Cr 4.3 ผู้ประเมินมีความสามารถในการจัดการการประเมินอย่างเหมาะสม	2	ผู้ประเมินสามารถเข้าใจบริบทของสถาบัน และเสนอ ทิศทางการพัฒนาสถาบันตามจุดเน้น เพื่อให้สามารถบรรลุเป้าหมายสูงสุดได้	ผู้ประเมินสามารถเข้าใจบริบทของสถาบัน และมีวิธีการที่จะเสนอแนวทางการพัฒนาอย่างเหมาะสม ในภาวะการณ์ปัจจุบัน	ผู้ประเมินเข้าใจบริบทของสถาบันไม่ชัดเจน ทำให้เสนอแนวทางการพัฒนา ไม่ สอดคล้องกับบริบทของสถาบัน	ผู้ประเมินยังไม่สามารถเข้าใจบริบท และมีมุมมอง ในการบริหารจัดการ และแนวทางการพัฒนาที่แคบ ยากต่อการพัฒนาสถาบัน			

มาตรฐาน และตัวบ่งชี้	น้ำหนัก (A)	ดีมาก (4)	ดี (3)	พอใช้ (2)	ปรับปรุง (1)	คะแนนการประเมิน (B)	ผลการประเมิน (AXB)	ข้อเสนอแนะ
Cr 4.4 กระบวนการประเมินที่คำนึงผลประโยชน์ทับซ้อนของผู้ประเมิน	2	การตัดสินใจผลการประเมินโปร่งใส ตรวจสอบได้ ไม่มี อิทธิพลเกี่ยวกับผลประโยชน์จากผู้ประเมินที่ส่งผลกระทบต่อประเมิน	การตัดสินใจผลการประเมินมีความถูกต้อง ชัดเจน และ ไม่ได้รับ อิทธิพลเกี่ยวกับผลประโยชน์จากผู้ประเมินที่ส่งผลกระทบต่อประเมิน	การตัดสินใจผลการประเมินมีความชัดเจน แต่ มี อิทธิพลเกี่ยวกับผลประโยชน์จากผู้ประเมินที่ส่งผลกระทบต่อประเมิน	การตัดสินใจผลการประเมินมีความคลาดเคลื่อนเนื่องจากได้รับอิทธิพลเกี่ยวกับผลประโยชน์ที่ขัดแย้งกับผู้ประเมิน			
ประสิทธิผลต้นทุน (Cost-effectiveness) 2 ตัวบ่งชี้	8							
Ce 5.1 การจัดการทรัพยากรการประเมินอย่างคุ้มค่า	4	การจัดการทรัพยากรการประเมินอย่างประหยัด คุ้มค่า บรรลุเป้าหมาย และผู้รับการประเมินมีความพึงพอใจ	การจัดการทรัพยากรการประเมินเป็นไปตามเป้าหมายที่วางไว้ อย่างเหมาะสม	มีการวางแผนการจัดการทรัพยากรการประเมิน แต่ไม่ระบุการใช้ทรัพยากร	ไม่มี การวางแผนการจัดการทรัพยากรการประเมิน			
Ce 5.2 การจัดสรรงบประมาณสำหรับการประเมินอย่างเหมาะสม	4	มีเอกสารแสดงถึงค่าใช้จ่ายในการประเมินอย่างสมเหตุสมผล จัดทำหลักฐานชัดเจน โปร่งใส สามารถตรวจสอบค่าใช้จ่าย ในการประเมินได้	มีการสรุปค่าใช้จ่ายที่ใช้ในการประเมิน มีหลักฐานแสดงการใช้จ่ายในบางส่วน แต่ยังไม่ชัดเจน	มีค่าใช้จ่ายในการประเมิน แต่ ไม่มี หลักฐานแสดงการใช้จ่าย ในการประเมิน	ไม่มี การวางแผนเกี่ยวกับงบประมาณในการประเมิน			
	100							

ข้อเสนอแนะในการนำผลการประเมินไปใช้ประโยชน์

.....

.....

.....

.....

.....



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รหัสรายงาน

แบบประเมินคุณภาพรายงานการประเมิน โดยผู้ทรงคุณวุฒิ

ชื่อมหาวิทยาลัย

คำชี้แจง โปรดทำเครื่องหมาย ลงในคะแนนผลการประเมินที่ตรงกับลักษณะรายงานที่ประเมิน

ข้อ	ลักษณะรายงานที่ประเมิน	ผลการประเมิน				
		0	1	2	3	4
1	การระบุข้อมูลพื้นฐานของสถานศึกษา	0	1	2	3	4
2	การเขียนค่านำของรายงานการประเมิน	0	1	2	3	4
3	การเก็บรวบรวมข้อมูลโดยภาพรวม	0	1	2	3	4
4	การเก็บรวบรวมข้อมูลตามองค์ประกอบและตัวบ่งชี้	0	1	2	3	4
5	การนำเสนอข้อมูลในรายงานการประเมิน	0	1	2	3	4
6	การประเมินตามเกณฑ์ของ สกอ.	0	1	2	3	4
7	ความตรงของการวิเคราะห์ผลการประเมิน	0	1	2	3	4
8	การนำเสนอผลการประเมิน	0	1	2	3	4
9	การสรุปผลการประเมิน	0	1	2	3	4
10	การให้ข้อเสนอแนะในการพัฒนาสถานศึกษา	0	1	2	3	4
11	ทิศทางการพัฒนาคุณภาพสถานศึกษาในอนาคต	0	1	2	3	4
12	รูปแบบรายงานผลการประเมินคุณภาพภายใน ระดับสถาบัน	0	1	2	3	4
13	ภาษาที่ใช้ในการเขียนรายงานผลการประเมินคุณภาพภายใน ระดับสถาบัน	0	1	2	3	4
14	ลักษณะการพิมพ์รายงานการประเมิน	0	1	2	3	4
15	ลักษณะการนำเสนอข้อมูลในภาคผนวก	0	1	2	3	4
16	ประโยชน์ของรายงานการประเมินในด้านการนำไปปฏิบัติ	0	1	2	3	4
17	ประโยชน์ของรายงานการประเมินในด้านวิชาการ	0	1	2	3	4
18	ประโยชน์ของรายงานการประเมินในด้านสังคม	0	1	2	3	4
19	ประสิทธิผลของรายงานการประเมิน	0	1	2	3	4
20	คุณภาพรายงานการประเมินโดยสรุปภาพรวม	0	1	2	3	4

จุฬาลงกรณ์มหาวิทยาลัย

เกณฑ์ประเมินคุณภาพสำหรับแบบประเมินคุณภาพรายงาน

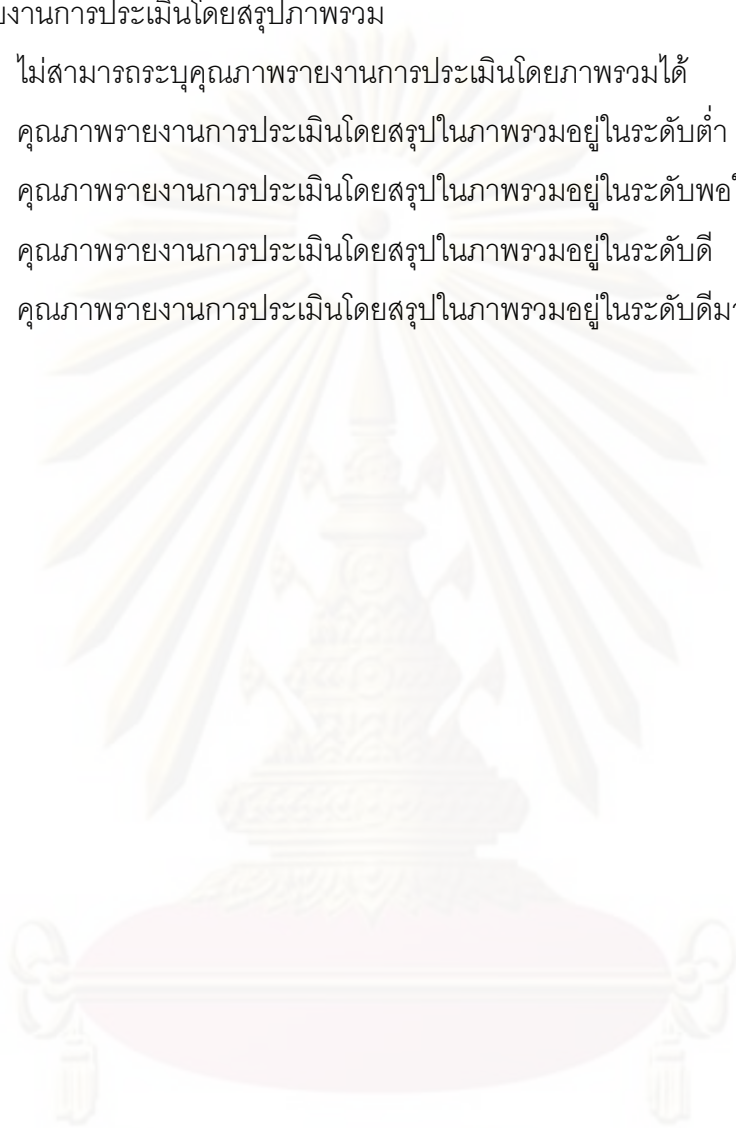
1. การระบุข้อมูลพื้นฐานของสถานศึกษา
 - 0 หมายถึง ไม่มีข้อมูลพื้นฐานของสถานศึกษา
 - 1 หมายถึง มีข้อมูลพื้นฐานของสถานศึกษาน้อยมาก ไม่ครอบคลุมหัวข้อที่กำหนด
 - 2 หมายถึง มีข้อมูลพื้นฐานของสถานศึกษาครบตามที่กำหนด แต่ไม่ครบสมบูรณ์
 - 3 หมายถึง มีข้อมูลพื้นฐานของสถานศึกษาครบตามที่กำหนด ครบสมบูรณ์ แต่ไม่เป็นปัจจุบัน
 - 4 หมายถึง มีข้อมูลพื้นฐานของสถานศึกษาครบตามที่กำหนด ครบสมบูรณ์ และเป็นปัจจุบัน
2. การเขียนคำนำของรายงานการประเมิน
 - 0 หมายถึง ไม่มีการเขียนคำนำ
 - 1 หมายถึง มีคำนำ แต่เขียนไม่ถูกต้อง
 - 2 หมายถึง มีคำนำ เขียนได้ถูกต้องชัดเจน ในระดับปานกลาง
 - 3 หมายถึง มีคำนำ เขียนได้ถูกต้องชัดเจน ในระดับดี
 - 4 หมายถึง มีคำนำ เขียนได้ถูกต้องชัดเจน ในระดับดีมาก
3. การเก็บรวบรวมข้อมูลโดยภาพรวม
 - 0 หมายถึง ไม่มีการเก็บรวบรวมข้อมูล
 - 1 หมายถึง มีกระบวนการเก็บรวบรวมข้อมูล แต่ยังไม่เป็นระบบ
 - 2 หมายถึง มีกระบวนการเก็บข้อมูลอย่างเป็นระบบชัดเจน
 - 3 หมายถึง ระบุและอธิบายวิธีการเก็บรวบรวมข้อมูลได้อย่างถูกต้อง เหมาะสมน่าเชื่อถือและมีความเป็นไปได้
 - 4 หมายถึง มีกระบวนการเก็บข้อมูลอย่างเป็นระบบชัดเจนและถูกต้อง เหมาะสมน่าเชื่อถือและมีความเป็นไปได้
4. การเก็บรวบรวมข้อมูลตามองค์ประกอบและตัวบ่งชี้
 - 0 หมายถึง ไม่ระบุวิธีการรวบรวมข้อมูล
 - 1 หมายถึง สามารถค้นหาเอกสาร/หลักฐานที่สถานศึกษาใช้ปฏิบัติงานตามปกติได้
 - 2 หมายถึง ใช้วิธีการสังเกตหรือสัมภาษณ์เพื่อรวบรวมข้อมูลได้อย่างเหมาะสม
 - 3 หมายถึง ใช้วิธีสังเกต สัมภาษณ์ และศึกษาเอกสารในการตรวจสอบข้อมูล
 - 4 หมายถึง ได้ข้อมูลถูกต้อง และครบถ้วนสำหรับประเมินตัวบ่งชี้หรือองค์ประกอบนั้นๆ
5. การนำเสนอข้อมูลในรายงานการประเมิน
 - 0 หมายถึง ไม่มีการนำเสนอข้อมูล
 - 1 หมายถึง มีข้อมูลไม่เพียงพอในการประเมิน
 - 2 หมายถึง มีข้อมูลเพียงพอในการประเมิน แต่นำเสนอไม่เป็นไปตามที่กำหนด

- 3 หมายถึง มีข้อมูลเพียงพอในการประเมิน แต่มีการนำเสนอ ข้อมูล สถิติ โครงการที่สำคัญ เพียงบางส่วน
- 4 หมายถึง มีข้อมูลที่ค้นพบอย่างชัดเจน เพียงพอสำหรับการประเมิน และมีการนำเสนอ ข้อมูล สถิติ โครงการ ที่สำคัญ พร้อมทั้งหลักฐานให้เห็นชัดเจน
6. การประเมินตามเกณฑ์ของ สกอ.
- 0 หมายถึง ไม่สามารถประเมินได้ตามตัวบ่งชี้ และเกณฑ์ สกอ.
- 1 หมายถึง สามารถประเมินได้ตามตัวบ่งชี้ และเกณฑ์ สกอ. จำนวนปานกลาง
- 2 หมายถึง สามารถประเมินได้ตามตัวบ่งชี้ และเกณฑ์ สกอ. จำนวนมาก
- 3 หมายถึง สามารถประเมินได้ถูกต้อง ครบถ้วนตามตัวบ่งชี้ และเกณฑ์ ของ สกอ.
- 4 หมายถึง สามารถประเมินได้ถูกต้อง ครบถ้วนตามเกณฑ์ของสกอ. และสถาบัน
7. ความตรงของการวิเคราะห์ผลการประเมิน
- 0 หมายถึง การวิเคราะห์ผลการประเมินไม่ถูกต้อง
- 1 หมายถึง การวิเคราะห์ผลการประเมินถูกต้อง เล็กน้อย
- 2 หมายถึง การวิเคราะห์ผลการประเมินถูกต้อง ปานกลาง
- 3 หมายถึง การวิเคราะห์ผลการประเมินถูกต้อง จำนวนมาก
- 4 หมายถึง การวิเคราะห์ผลการประเมินถูกต้อง ชัดเจน และเชื่อถือ
8. การนำเสนอผลการประเมิน
- 0 หมายถึง ไม่นำเสนอผลการประเมินรายองค์ประกอบ ตัวบ่งชี้
- 1 หมายถึง มีผลการประเมินในแต่ละองค์ประกอบครอบคลุมทุกตัวบ่งชี้
- 2 หมายถึง ระบุจุดเด่น จุดที่ควรพัฒนา และปัจจัยที่ทำให้เกิดจุดเด่นและจุดที่ควรพัฒนาของ สถานศึกษาแต่ละด้านได้ชัดเจน
- 3 หมายถึง ข้อเสนอแนะมีผลการประเมินและเหตุผลสนับสนุนอย่างเพียงพอ สามารถนำไป ปฏิบัติได้
- 4 หมายถึง ข้อเสนอแนะมีผลการประเมินและเหตุผลสนับสนุนอย่างเพียงพอ และมีการเรียบเรียง การนำเสนอผลถูกต้องตามหลักการรายงานผลการประเมิน
9. การสรุปผลการประเมิน
- 0 หมายถึง ไม่มีการสรุปผลการประเมิน
- 1 หมายถึง มีการสรุปผลการประเมินตามองค์ประกอบ
- 2 หมายถึง มีการสรุปผลการประเมินตามองค์ประกอบ และตามมาตรฐานการอุดมศึกษา
- 3 หมายถึง มีการสรุปผลการประเมินตามองค์ประกอบ มาตรฐานการอุดมศึกษาและมุมมองด้าน การบริหารจัดการ

- 4 หมายถึง มีการสรุปผลการประเมินถูกต้อง ครบถ้วน และชัดเจน
10. การให้ข้อเสนอแนะในการพัฒนาสถานศึกษา
- 0 หมายถึง ไม่ให้ข้อเสนอแนะ
- 1 หมายถึง ให้ข้อเสนอแนะ และระบุจุดที่ควรพัฒนา
- 2 หมายถึง ให้ข้อเสนอแนะและจุดที่ควรพัฒนาสอดคล้องกัน
- 3 หมายถึง ให้ข้อเสนอแนะสอดคล้องกับจุดที่ควรพัฒนา และสามารถปฏิบัติได้ทั้งหมด
- 4 หมายถึง สรุปให้เห็นภาพรวมของข้อเสนอแนะในการพัฒนาสถาบัน
11. ทิศทางการพัฒนาคุณภาพสถานศึกษาในอนาคต
- 0 หมายถึง ไม่ระบุทิศทางในการพัฒนา
- 1 หมายถึง ระบุทิศทางการพัฒนาไม่สอดคล้องกับบริบทของสถานศึกษา
- 2 หมายถึง ระบุทิศทางการพัฒนาสอดคล้องกับบริบทของสถานศึกษา แต่ยังไม่ปฏิบัติไม่ได้
- 3 หมายถึง ระบุทิศทางการพัฒนาสอดคล้องกับบริบทของสถานศึกษา และสามารถปฏิบัติได้
- 4 หมายถึง ระบุทิศทางการพัฒนาสอดคล้องกับบริบทของสถานศึกษา และเป็นไปตามนโยบายสกอ.
12. รูปแบบรายงานการประเมินคุณภาพภายในสถานศึกษา
- 0 หมายถึง ไม่มีหัวข้อตรงกับรูปแบบที่ สกอ. กำหนด
- 1 หมายถึง มีหัวข้อบางส่วนตรงกับรูปแบบที่ สกอ. กำหนด
- 2 หมายถึง มีหัวข้อตรงตามรูปแบบที่ สกอ. กำหนด แต่เรียงลำดับไม่ถูกต้อง
- 3 หมายถึง มีหัวข้อตรงตามรูปแบบที่ สกอ. กำหนด เรียงลำดับถูกต้อง
- 4 หมายถึง มีหัวข้อตรงตามรูปแบบที่ สกอ. กำหนด เนื้อหาชัดเจน และมีคุณค่าทางวิชาการในตัวเอง
13. ภาษาที่ใช้ในการเขียนรายงานการประเมินคุณภาพภายในสถานศึกษา
- 0 หมายถึง ใช้ภาษาไม่เหมาะสม
- 1 หมายถึง ใช้ภาษาเขียนเหมาะสม
- 2 หมายถึง ใช้ภาษาเหมาะสม สื่อความได้
- 3 หมายถึง ใช้ภาษาเชิงวิชาการ มีสาระเหมาะสม สื่อความได้
- 4 หมายถึง ใช้ภาษาเชิงวิชาการ มีสาระเหมาะสม สื่อความได้ และสอดคล้องเชื่อมโยงกันตลอด
14. ลักษณะการพิมพ์รายงานการประเมิน
- 0 หมายถึง มีคำผิด และตัดคำไม่ถูกต้อง
- 1 หมายถึง การสะกดคำถูกต้อง
- 2 หมายถึง การสะกดคำถูกต้อง และการตัดคำถูกต้องเหมาะสม

- 3 หมายถึง การสะกดคำ การตัดคำและการเว้นวรรคถูกต้องเหมาะสม
- 4 หมายถึง การสะกดคำ การตัดคำและการเว้นวรรคถูกต้องเหมาะสม ย่อหน้าและแบ่ง
ความสำคัญอย่างเหมาะสม
15. ลักษณะการนำเสนอข้อมูลในภาคผนวก
- 0 หมายถึง ไม่มีภาคผนวก
- 1 หมายถึง มีภาคผนวก แต่มีเฉพาะกำหนดการประเมิน
- 2 หมายถึง มีภาคผนวก มีกำหนดการประเมิน ข้อมูลการสัมภาษณ์
- 3 หมายถึง มีภาคผนวก มีกำหนดการประเมิน สรุปข้อมูลการสัมภาษณ์ และการตรวจเยี่ยม
- 4 หมายถึง มีภาคผนวก มีกำหนดการประเมิน สรุปข้อมูลการสัมภาษณ์ การตรวจเยี่ยม และ
รายละเอียดการวิเคราะห์ข้อมูล
16. ประโยชน์ของรายงานการประเมินในด้านการนำไปปฏิบัติ
- 0 หมายถึง ประโยชน์ของรายงาน ยังไม่ชัดเจนในการนำไปปฏิบัติ
- 1 หมายถึง เห็นประโยชน์ชัดเจน เป็นรูปธรรม
- 2 หมายถึง มีการจัดหมวดหมู่ ประโยชน์ที่ควรจะได้รับ
- 3 หมายถึง มีการสรุปภาพรวมของสิ่งที่สถาบันควรดำเนินการต่อไป
- 4 หมายถึง มีการคาดหมายถึงผลที่จะเกิดขึ้น ถ้าได้ดำเนินการตามที่เสนอแนะ
17. ประโยชน์ของรายงานการประเมินในด้านวิชาการ
- 0 หมายถึง ไม่เห็นประโยชน์ในด้านวิชาการ
- 1 หมายถึง มีประโยชน์ในการนำไปทบทวนการดำเนินการของสถาบัน
- 2 หมายถึง มีประโยชน์นำไปปรับปรุงตนเองในการบริหารจัดการภายในสถาบัน
- 3 หมายถึง มีประโยชน์นำไปพัฒนาระบบประกันคุณภาพของสถาบัน
- 4 หมายถึง มีประโยชน์นำไปวางแผนพัฒนาสถาบันสู่ความเป็นเลิศ
18. ประโยชน์ของรายงานการประเมินในด้านสังคม
- 0 หมายถึง ไม่ประโยชน์ในด้านสังคม
- 1 หมายถึง มีประโยชน์ต่อสังคม ระดับสถาบัน
- 2 หมายถึง มีประโยชน์ต่อสังคม ระดับพื้นที่
- 3 หมายถึง มีประโยชน์ต่อสังคม ระดับภูมิภาค
- 4 หมายถึง มีประโยชน์ด้านสังคม ระดับชาติ
19. ประสิทธิภาพของรายงานการประเมิน
- 0 หมายถึง ไม่เป็นไปตามวัตถุประสงค์การประเมิน
- 1 หมายถึง เป็นไปตามวัตถุประสงค์การประเมินน้อย

- 2 หมายถึง เป็นไปตามวัตถุประสงค์การประเมิน บางส่วน
- 3 หมายถึง เป็นไปตามวัตถุประสงค์การประเมิน ส่วนมาก
- 4 หมายถึง เป็นไปตามวัตถุประสงค์การประเมินครบถ้วน
20. คุณภาพรายงานการประเมินโดยสรุปภาพรวม
- 0 หมายถึง ไม่สามารถระบุคุณภาพรายงานการประเมินโดยภาพรวมได้
- 1 หมายถึง คุณภาพรายงานการประเมินโดยสรุปในภาพรวมอยู่ในระดับต่ำ
- 2 หมายถึง คุณภาพรายงานการประเมินโดยสรุปในภาพรวมอยู่ในระดับพอใช้
- 3 หมายถึง คุณภาพรายงานการประเมินโดยสรุปในภาพรวมอยู่ในระดับดี
- 4 หมายถึง คุณภาพรายงานการประเมินโดยสรุปในภาพรวมอยู่ในระดับดีมาก



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ค
ภาพกิจกรรมการอบรมนักประเมินอภิมาน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ค ภาพกิจกรรมการจัดอบรมนักประเมินอภิมาน

รุ่นที่ 1 จัดขึ้นระหว่างวันที่ 19-20 กันยายน 2552 เวลา 08.30-16.30 น. ณ ห้อง 605และ504

อาคารประชุมสุข อาชีวอารุง (อาคาร 3) คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย



ภาพ 1 บรรยายแนวคิด หลักการประเมินอภิมาน มาตรฐาน เกณฑ์การประเมินอภิมาน



ภาพ 2 ศึกษามาตรฐาน ตัวบ่งชี้ เกณฑ์การประเมินอภิมานแต่ละมิติ



ภาพ 3 วิทยากรรับเชิญเล่าประสบการณ์การประเมินคุณภาพภายในและการเขียนรายงาน



ภาพ 4 ฝึกปฏิบัติประเมินอภิมานตามมาตรฐาน และเกณฑ์การประเมินอภิมาน

รุ่นที่ 2 จัดขึ้นระหว่าง วันจันทร์ที่ 4 และวันอังคารที่ 5 มกราคม 2553 เวลา 08.30-16.30 น.
ณ ห้องบงกช อาคาร 10 ชั้น 6 มหาวิทยาลัยศรีปทุม



ภาพ 1 พิธีเปิดการฝึกอบรมนักประเมินอภิมาน โดยรองอธิการบดี2



ภาพ 2 บรรยายหลักการและแนวคิดการประเมินอภิมาน



ภาพ 3 ศึกษามาตรฐาน ตัวบ่งชี้ เกณฑ์การประเมินอภิมานแต่ละมิติ



ภาพ 4 วิทยากรเล่าประสบการณ์ เทคนิคการประเมินคุณภาพภายในและการเขียนรายงาน



ภาคผนวก ง
Print Out การวิเคราะห์ข้อมูล

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาคผนวก ง Print Out การวิเคราะห์ข้อมูล

DATE: 4/ 1/2010

TIME: 23:07

L I S R E L 8.72

BY

Karl G. Joreskog & Dag Sörbom

This program is published exclusively by
 Scientific Software International, Inc.
 7383 N. Lincoln Avenue, Suite 100
 Lincolnwood, IL 60712, U.S.A.
 Phone: (800)247-6113, (847)675-0720, Fax: (847)675-2140
 Copyright by Scientific Software International, Inc., 1981-2005
 Use of this program is subject to the terms specified in the
 Universal Copyright Convention.
 Website: www.ssicentral.com

The following lines were read from file D:\proposal\วิเคราะห์ข้อมูล\Lisrel26 March\VALIDITY.spl:

! MEASUREMENT OF VALIDITY MODEL

DA NI=13 NO=50
 LA FI=LA01.DAT
 KM FI=KM01.DAT
 SD FI=SD01.DAT
 MO NX=13 NK=1 LX=FR TD=SY
 FR TD(9,10) TD(4,1) TD(10,7)
 LK
 VALIDITY
 PD
 OU MI FS

! MEASUREMENT OF VALIDITY MODEL

Number of Input Variables 13
 Number of Y - Variables 0
 Number of X - Variables 13
 Number of ETA - Variables 0
 Number of KSI - Variables 1
 Number of Observations 50

! MEASUREMENT OF VALIDITY MODEL

Covariance Matrix

	Val.1	Val.2	Val.3	Val.4	Val.5	Val.6
Val.1	0.64					
Val.2	0.19	0.82				
Val.3	0.34	0.33	0.98			
Val.4	0.58	0.25	0.60	1.62		
Val.5	0.20	0.15	0.33	0.40	1.01	
Val.6	0.22	0.03	0.13	0.18	0.23	0.63
Val.7	0.30	0.35	0.42	0.25	0.43	0.18
Val.8	0.24	0.11	0.23	0.43	0.23	0.22
Val.9	0.32	0.18	0.25	0.26	0.40	0.23
Val.10	0.30	0.09	0.24	0.29	0.37	0.12
Val.11	0.14	0.14	0.17	0.32	0.19	0.23
Val.12	0.19	0.31	0.14	0.28	0.31	0.13
Val.13	0.41	0.32	0.37	0.50	0.62	0.44

Covariance Matrix

	Val.7	Val.8	Val.9	Val.10	Val.11	Val.12
Val.7	0.81					
Val.8	0.28	0.47				
Val.9	0.28	0.24	0.99			
Val.10	0.09	0.17	0.61	0.77		
Val.11	0.12	0.18	0.26	0.20	0.34	
Val.12	0.17	0.23	0.30	0.27	0.24	0.91
Val.13	0.37	0.28	0.33	0.39	0.31	0.39

Covariance Matrix

	Val.13
Val.13	1.32

! MEASUREMENT OF VALIDITY MODEL

Parameter Specifications

LAMBDA-X

	VALIDITY
Val.1	1
Val.2	2
Val.3	3
Val.4	4
Val.5	5
Val.6	6
Val.7	7
Val.8	8
Val.9	9
Val.10	10
Val.11	11
Val.12	12
Val.13	13

THETA-DELTA

	Val.1	Val.2	Val.3	Val.4	Val.5	Val.6
Val.1	14					
Val.2	0	15				
Val.3	0	0	16			
Val.4	17	0	0	18		
Val.5	0	0	0	0	19	
Val.6	0	0	0	0	0	20
Val.7	0	0	0	0	0	0
Val.8	0	0	0	0	0	0
Val.9	0	0	0	0	0	0
Val.10	0	0	0	0	0	0
Val.11	0	0	0	0	0	0
Val.12	0	0	0	0	0	0
Val.13	0	0	0	0	0	0

THETA-DELTA

	Val.7	Val.8	Val.9	Val.10	Val.11	Val.12
Val.7	21					
Val.8	0	22				
Val.9	0	0	23			
Val.10	24	0	25	26		
Val.11	0	0	0	0	27	
Val.12	0	0	0	0	0	28

Val.13 0 0 0 0 0 0

THETA-DELTA

 Val.13

Val.13 29

! MEASUREMENT OF VALIDITY MODEL

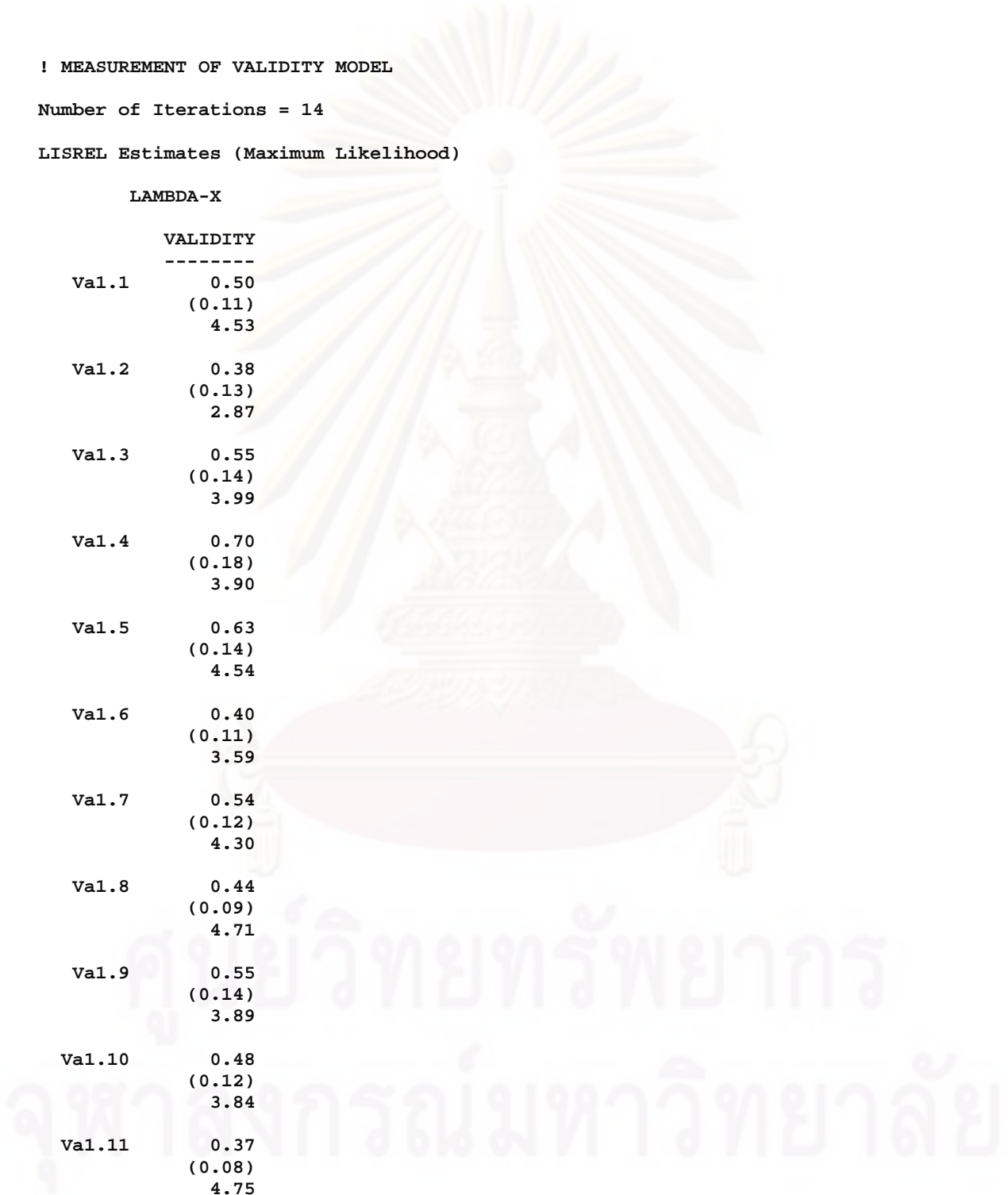
Number of Iterations = 14

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	VALIDITY

Val.1	0.50 (0.11) 4.53
Val.2	0.38 (0.13) 2.87
Val.3	0.55 (0.14) 3.99
Val.4	0.70 (0.18) 3.90
Val.5	0.63 (0.14) 4.54
Val.6	0.40 (0.11) 3.59
Val.7	0.54 (0.12) 4.30
Val.8	0.44 (0.09) 4.71
Val.9	0.55 (0.14) 3.89
Val.10	0.48 (0.12) 3.84
Val.11	0.37 (0.08) 4.75
Val.12	0.47 (0.14) 3.44



Val.13 0.80
 (0.15)
 5.23

PHI

VALIDITY

 1.00

THETA-DELTA

	Val.1	Val.2	Val.3	Val.4	Val.5	Val.6
	-----	-----	-----	-----	-----	-----
Val.1	0.40 (0.09) 4.45					
Val.2	- -	0.68 (0.14) 4.79				
Val.3	- -	- -	0.68 (0.15) 4.60			
Val.4	0.23 (0.11) 2.05	- -	- -	1.13 (0.25) 4.60		
Val.5	- -	- -	- -	- -	0.62 (0.14) 4.47	
Val.6	- -	- -	- -	- -	- -	0.46 (0.10) 4.68
Val.7	- -	- -	- -	- -	- -	- -
Val.8	- -	- -	- -	- -	- -	- -
Val.9	- -	- -	- -	- -	- -	- -
Val.10	- -	- -	- -	- -	- -	- -
Val.11	- -	- -	- -	- -	- -	- -
Val.12	- -	- -	- -	- -	- -	- -
Val.13	- -	- -	- -	- -	- -	- -

THETA-DELTA

	Val.7	Val.8	Val.9	Val.10	Val.11	Val.12
	-----	-----	-----	-----	-----	-----
Val.7	0.52 (0.12) 4.50					
Val.8	- -	0.28 (0.06) 4.42				
Val.9	- -	- -	0.69			

			(0.15)			
			4.60			
Val.10	-0.16	- -	0.34	0.54		
	(0.07)		(0.11)	(0.12)		
	-2.24		3.12	4.62		
Val.11	- -	- -	- -	- -	0.20	
					(0.04)	
					4.41	
Val.12	- -	- -	- -	- -	- -	0.69
						(0.15)
						4.71
Val.13	- -	- -	- -	- -	- -	- -

THETA-DELTA

	Val.13

Val.13	0.69
	(0.16)
	4.24

Squared Multiple Correlations for X - Variables

Val.1	Val.2	Val.3	Val.4	Val.5	Val.6
-----	-----	-----	-----	-----	-----
0.39	0.18	0.31	0.30	0.39	0.26

Squared Multiple Correlations for X - Variables

Val.7	Val.8	Val.9	Val.10	Val.11	Val.12
-----	-----	-----	-----	-----	-----
0.36	0.41	0.30	0.30	0.41	0.24

Squared Multiple Correlations for X - Variables

Val.13

0.48

Goodness of Fit Statistics

Degrees of Freedom = 62
 Minimum Fit Function Chi-Square = 77.78 (P = 0.085)
 Normal Theory Weighted Least Squares Chi-Square = 62.87 (P = 0.45)
 Estimated Non-centrality Parameter (NCP) = 0.87
 90 Percent Confidence Interval for NCP = (0.0 ; 23.75)

Minimum Fit Function Value = 1.59
 Population Discrepancy Function Value (F0) = 0.018
 90 Percent Confidence Interval for F0 = (0.0 ; 0.48)
 Root Mean Square Error of Approximation (RMSEA) = 0.017
 90 Percent Confidence Interval for RMSEA = (0.0 ; 0.088)
 P-Value for Test of Close Fit (RMSEA < 0.05) = 0.69

Expected Cross-Validation Index (ECVI) = 2.47
 90 Percent Confidence Interval for ECVI = (2.45 ; 2.93)
 ECVI for Saturated Model = 3.71
 ECVI for Independence Model = 10.18

Chi-Square for Independence Model with 78 Degrees of Freedom = 472.84
 Independence AIC = 498.84

Model AIC = 120.87
 Saturated AIC = 182.00
 Independence CAIC = 536.70
 Model CAIC = 205.32
 Saturated CAIC = 446.99

Normed Fit Index (NFI) = 0.84
 Non-Normed Fit Index (NNFI) = 0.95
 Parsimony Normed Fit Index (PNFI) = 0.66
 Comparative Fit Index (CFI) = 0.96
 Incremental Fit Index (IFI) = 0.96
 Relative Fit Index (RFI) = 0.79

Critical N (CN) = 58.20

Root Mean Square Residual (RMR) = 0.066
 Standardized RMR = 0.078
 Goodness of Fit Index (GFI) = 0.84
 Adjusted Goodness of Fit Index (AGFI) = 0.76
 Parsimony Goodness of Fit Index (PGFI) = 0.57

! MEASUREMENT OF VALIDITY MODEL

Modification Indices and Expected Change

No Non-Zero Modification Indices for LAMBDA-X

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

	Val.1	Val.2	Val.3	Val.4	Val.5	Val.6
Val.1	- -					
Val.2	0.00	- -				
Val.3	0.10	1.70	- -			
Val.4	- -	0.01	2.44	- -		
Val.5	2.77	1.10	0.04	0.04	- -	
Val.6	0.61	2.63	1.69	1.60	0.09	- -
Val.7	1.86	2.51	3.12	2.60	2.71	1.32
Val.8	0.05	0.85	0.06	2.50	0.63	0.94
Val.9	0.06	0.01	0.56	0.55	0.01	1.06
Val.10	1.21	0.33	0.32	0.31	1.33	2.86
Val.11	2.73	0.00	0.54	2.26	1.11	4.22
Val.12	0.34	1.97	1.61	0.03	0.03	0.67
Val.13	0.10	0.02	0.65	0.30	2.29	2.87

Modification Indices for THETA-DELTA

	Val.7	Val.8	Val.9	Val.10	Val.11	Val.12
Val.7	- -					
Val.8	0.33	- -				
Val.9	0.03	0.07	- -			
Val.10	- -	0.34	- -	- -		
Val.11	4.59	0.24	1.80	0.57	- -	
Val.12	1.16	0.17	0.20	0.01	2.04	- -
Val.13	0.33	1.56	1.75	0.44	0.13	0.04

Modification Indices for THETA-DELTA

	Val.13
Val.13	- -

Expected Change for THETA-DELTA

	Val.1	Val.2	Val.3	Val.4	Val.5	Val.6
Val.1	- -					
Val.2	0.00	- -				
Val.3	0.02	0.13	- -			
Val.4	- -	-0.01	0.19	- -		
Val.5	-0.12	-0.10	-0.02	0.02	- -	
Val.6	0.05	-0.14	-0.11	-0.13	-0.03	- -
Val.7	0.09	0.14	0.16	-0.17	0.15	-0.08
Val.8	-0.01	-0.06	-0.02	0.13	-0.05	0.05
Val.9	0.01	0.01	-0.06	-0.07	-0.01	0.07
Val.10	0.06	-0.04	0.04	-0.05	0.08	-0.10
Val.11	-0.07	0.00	-0.04	0.10	-0.06	0.10
Val.12	-0.04	0.14	-0.13	-0.02	0.02	-0.07
Val.13	0.03	0.02	-0.09	-0.07	0.17	0.15

Expected Change for THETA-DELTA

	Val.7	Val.8	Val.9	Val.10	Val.11	Val.12
Val.7	- -					
Val.8	0.03	- -				
Val.9	-0.02	0.01	- -			
Val.10	- -	-0.03	- -	- -		
Val.11	-0.11	0.02	0.06	-0.03	- -	
Val.12	-0.10	0.03	0.04	-0.01	0.08	- -
Val.13	-0.06	-0.09	-0.12	0.05	0.02	0.02

Expected Change for THETA-DELTA

Val.13	
Val.13	- -

Maximum Modification Index is 4.59 for Element (11, 7) of THETA-DELTA

! MEASUREMENT OF VALIDITY MODEL

Factor Scores Regressions

KSI

	Val.1	Val.2	Val.3	Val.4	Val.5	Val.6
VALIDITY	0.14	0.07	0.11	0.05	0.13	0.12

KSI

	Val.7	Val.8	Val.9	Val.10	Val.11	Val.12
VALIDITY	0.18	0.21	0.03	0.15	0.25	0.09

KSI

Val.13	
VALIDITY	0.15

Time used: 0.031 Seconds

```
! MEASUREMENT OF UTILITY MODEL
DA NI=10 NO=50
LA FI=LA02.DAT
KM FI=KM02.DAT
SD FI=SD02.DAT
MO NX=10 NK=1 LX=FR TD=SY
FR TD(7,2)TD(9,1)TD(9,3)TD(10,9)
FR TD(6,3)
LK
UTILITY
PD
OU MI FS
```

```
! MEASUREMENT OF UTILITY MODEL
```

```
Number of Input Variables 10
Number of Y - Variables 0
Number of X - Variables 10
Number of ETA - Variables 0
Number of KSI - Variables 1
Number of Observations 50
```

```
! MEASUREMENT OF UTILITY MODEL
```

Covariance Matrix

	Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6
Ut2.1	1.08					
Ut2.2	0.30	0.93				
Ut2.3	0.15	0.13	0.33			
Ut2.4	0.24	0.24	0.06	0.45		
Ut2.5	0.41	0.43	0.12	0.43	1.31	
Ut2.6	0.22	0.24	-0.03	0.32	0.61	0.87
Ut2.7	0.33	0.04	0.05	0.16	0.31	0.39
Ut2.8	0.14	0.35	0.08	0.26	0.29	0.11
Ut2.9	0.42	0.29	-0.03	0.17	0.30	0.22
Ut2.10	0.34	0.27	0.11	0.15	0.34	0.21

Covariance Matrix

	Ut2.7	Ut2.8	Ut2.9	Ut2.10
Ut2.7	0.52			
Ut2.8	0.08	0.72		
Ut2.9	0.19	0.14	0.46	
Ut2.10	0.25	0.11	0.31	0.62

```
! MEASUREMENT OF UTILITY MODEL
```

Parameter Specifications

LAMBDA-X

	UTILITY
Ut2.1	1
Ut2.2	2
Ut2.3	3
Ut2.4	4
Ut2.5	5
Ut2.6	6
Ut2.7	7
Ut2.8	8
Ut2.9	9
Ut2.10	10

THETA-DELTA

	Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6
Ut2.1	11					
Ut2.2	0	12				
Ut2.3	0	0	13			
Ut2.4	0	0	0	14		
Ut2.5	0	0	0	0	15	
Ut2.6	0	0	16	0	0	17
Ut2.7	0	18	0	0	0	0
Ut2.8	0	0	0	0	0	0
Ut2.9	21	0	22	0	0	0
Ut2.10	0	0	0	0	0	0

THETA-DELTA

	Ut2.7	Ut2.8	Ut2.9	Ut2.10
Ut2.7	19			
Ut2.8	0	20		
Ut2.9	0	0	23	
Ut2.10	0	0	24	25

! MEASUREMENT OF UTILITY MODEL

Number of Iterations = 17

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	UTILITY
Ut2.1	0.56 (0.15) 3.84
Ut2.2	0.56 (0.14) 4.11
Ut2.3	0.18 (0.09) 2.04
Ut2.4	0.45 (0.09) 4.98
Ut2.5	0.81 (0.15) 5.41
Ut2.6	0.64 (0.13) 5.08
Ut2.7	0.45 (0.10) 4.55
Ut2.8	0.35 (0.12) 2.86

Ut2.9 0.42
 (0.09)
 4.47

Ut2.10 0.45
 (0.11)
 4.04

PHI

UTILITY

 1.00

THETA-DELTA

	Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6
	-----	-----	-----	-----	-----	-----
Ut2.1	0.76 (0.17) 4.63					
Ut2.2	- -	0.61 (0.14) 4.49				
Ut2.3	- -	- -	0.30 (0.06) 4.93			
Ut2.4	- -	- -	- -	0.25 (0.06) 4.40		
Ut2.5	- -	- -	- -	- -	0.65 (0.15) 4.24	
Ut2.6	- -	- -	-0.14 (0.05) -2.54	- -	- -	0.47 (0.11) 4.23
Ut2.7	- -	-0.21 (0.07) -2.84	- -	- -	- -	- -
Ut2.8	- -	- -	- -	- -	- -	- -
Ut2.9	0.17 (0.07) 2.39	- -	-0.12 (0.04) -3.07	- -	- -	- -
Ut2.10	- -	- -	- -	- -	- -	- -

THETA-DELTA

	Ut2.7	Ut2.8	Ut2.9	Ut2.10
	-----	-----	-----	-----
Ut2.7	0.31 (0.07) 4.37			
Ut2.8	- -	0.60 (0.12) 4.81		

Ut2.9	- -	- -	0.28 (0.06) 4.67	
Ut2.10	- -	- -	0.11 (0.05) 2.10	0.43 (0.09) 4.59

Squared Multiple Correlations for X - Variables

Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6
-----	-----	-----	-----	-----	-----
0.29	0.34	0.10	0.44	0.50	0.46

Squared Multiple Correlations for X - Variables

Ut2.7	Ut2.8	Ut2.9	Ut2.10
-----	-----	-----	-----
0.39	0.17	0.38	0.32

Goodness of Fit Statistics

Degrees of Freedom = 30
 Minimum Fit Function Chi-Square = 34.35 (P = 0.27)
 Normal Theory Weighted Least Squares Chi-Square = 31.62 (P = 0.39)
 Estimated Non-centrality Parameter (NCP) = 1.62
 90 Percent Confidence Interval for NCP = (0.0 ; 19.32)

Minimum Fit Function Value = 0.70
 Population Discrepancy Function Value (F0) = 0.033
 90 Percent Confidence Interval for F0 = (0.0 ; 0.39)
 Root Mean Square Error of Approximation (RMSEA) = 0.033
 90 Percent Confidence Interval for RMSEA = (0.0 ; 0.11)
 P-Value for Test of Close Fit (RMSEA < 0.05) = 0.56

Expected Cross-Validation Index (ECVI) = 1.67
 90 Percent Confidence Interval for ECVI = (1.63 ; 2.03)
 ECVI for Saturated Model = 2.24
 ECVI for Independence Model = 5.99

Chi-Square for Independence Model with 45 Degrees of Freedom = 273.62
 Independence AIC = 293.62
 Model AIC = 81.62
 Saturated AIC = 110.00
 Independence CAIC = 322.74
 Model CAIC = 154.42
 Saturated CAIC = 270.16

Normed Fit Index (NFI) = 0.87
 Non-Normed Fit Index (NNFI) = 0.97
 Parsimony Normed Fit Index (PNFI) = 0.58
 Comparative Fit Index (CFI) = 0.98
 Incremental Fit Index (IFI) = 0.98
 Relative Fit Index (RFI) = 0.81

Critical N (CN) = 73.61

Root Mean Square Residual (RMR) = 0.055
 Standardized RMR = 0.074
 Goodness of Fit Index (GFI) = 0.89
 Adjusted Goodness of Fit Index (AGFI) = 0.79
 Parsimony Goodness of Fit Index (PGFI) = 0.48

! MEASUREMENT OF UTILITY MODEL

Modification Indices and Expected Change

No Non-Zero Modification Indices for LAMBDA-X

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

	Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6
Ut2.1	- -					
Ut2.2	0.00	- -				
Ut2.3	0.06	0.21	- -			
Ut2.4	0.00	0.93	0.06	- -		
Ut2.5	0.08	0.90	0.01	1.95	- -	
Ut2.6	2.38	0.74	- -	0.77	2.36	- -
Ut2.7	2.05	- -	0.08	2.76	2.30	2.70
Ut2.8	0.49	2.30	0.03	3.91	0.01	2.98
Ut2.9	- -	1.27	- -	0.06	0.17	0.00
Ut2.10	1.48	0.25	0.04	0.86	0.03	1.37

Modification Indices for THETA-DELTA

	Ut2.7	Ut2.8	Ut2.9	Ut2.10
Ut2.7	- -			
Ut2.8	0.31	- -		
Ut2.9	0.22	0.07	- -	
Ut2.10	2.35	0.56	- -	- -

Expected Change for THETA-DELTA

	Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6
Ut2.1	- -					
Ut2.2	0.00	- -				
Ut2.3	0.02	0.03	- -			
Ut2.4	0.00	-0.06	-0.01	- -		
Ut2.5	-0.03	-0.10	0.00	0.10	- -	
Ut2.6	-0.13	-0.08	- -	0.05	0.15	- -
Ut2.7	0.10	- -	-0.01	-0.08	-0.12	0.11
Ut2.8	-0.06	0.13	-0.01	0.12	0.01	-0.14
Ut2.9	- -	0.06	- -	-0.01	-0.02	0.00
Ut2.10	0.11	0.04	0.01	-0.04	-0.01	-0.08

Expected Change for THETA-DELTA

	Ut2.7	Ut2.8	Ut2.9	Ut2.10
Ut2.7	- -			
Ut2.8	-0.04	- -		
Ut2.9	-0.02	0.01	- -	
Ut2.10	0.08	-0.05	- -	- -

Maximum Modification Index is 3.91 for Element (8, 4) of THETA-DELTA

! MEASUREMENT OF UTILITY MODEL

Factor Scores Regressions

KSI

	Ut2.1	Ut2.2	Ut2.3	Ut2.4	Ut2.5	Ut2.6
UTILITY	0.03	0.22	0.30	0.21	0.15	0.25

KSI

	Ut2.7	Ut2.8	Ut2.9	Ut2.10
UTILITY	0.32	0.07	0.27	0.06

Time used: 0.063 Seconds

! MEASUREMENT OF ETHICALITY MODEL

DA NI=9 NO=50
 LA FI=LA03.DAT
 KM FI=KM03.DAT
 SD FI=SD03.DAT
 MO NX=9 NK=1 LX=FR TD=SY
 FR TD(7,2) TD(5,4) TD(8,7)
 LK
 ETHICALITY
 PD
 OU MI FS

! MEASUREMENT OF ETHICALITY MODEL

Number of Input Variables 9
 Number of Y - Variables 0
 Number of X - Variables 9
 Number of ETA - Variables 0
 Number of KSI - Variables 1
 Number of Observations 50

! MEASUREMENT OF ETHICALITY MODEL

Covariance Matrix

	Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6
Et3.1	0.41					
Et3.2	0.27	0.66				
Et3.3	0.20	0.22	0.52			
Et3.4	0.22	0.24	0.19	0.66		
Et3.5	0.15	0.36	0.11	0.32	0.61	
Et3.6	0.28	0.37	0.24	0.30	0.35	0.62
Et3.7	0.22	0.47	0.20	0.19	0.29	0.29
Et3.8	0.18	0.43	0.14	0.12	0.30	0.33
Et3.9	0.27	0.39	0.21	0.19	0.23	0.41

Covariance Matrix

	Et3.7	Et3.8	Et3.9
Et3.7	0.66		
Et3.8	0.46	0.83	
Et3.9	0.27	0.42	0.77

! MEASUREMENT OF ETHICALITY MODEL

Parameter Specifications

LAMBDA-X

ETHICALI	
Et3.1	1
Et3.2	2
Et3.3	3
Et3.4	4
Et3.5	5
Et3.6	6

Et3.7 7
 Et3.8 8
 Et3.9 9

THETA-DELTA

	Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6
Et3.1	10					
Et3.2	0	11				
Et3.3	0	0	12			
Et3.4	0	0	0	13		
Et3.5	0	0	0	14	15	
Et3.6	0	0	0	0	0	16
Et3.7	0	17	0	0	0	0
Et3.8	0	0	0	0	0	0
Et3.9	0	0	0	0	0	0

THETA-DELTA

	Et3.7	Et3.8	Et3.9
Et3.7	18		
Et3.8	19	20	
Et3.9	0	0	21

! MEASUREMENT OF ETHICALITY MODEL

Number of Iterations = 8

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	ETHICALI
Et3.1	0.42 (0.09) 4.87
Et3.2	0.64 (0.10) 6.23
Et3.3	0.35 (0.10) 3.43
Et3.4	0.40 (0.12) 3.46
Et3.5	0.49 (0.11) 4.61
Et3.6	0.63 (0.10) 6.43
Et3.7	0.51 (0.11) 4.63
Et3.8	0.57 (0.12)

4.60
 Et3.9 0.62
 (0.12)
 5.32

PHI

ETHICALI

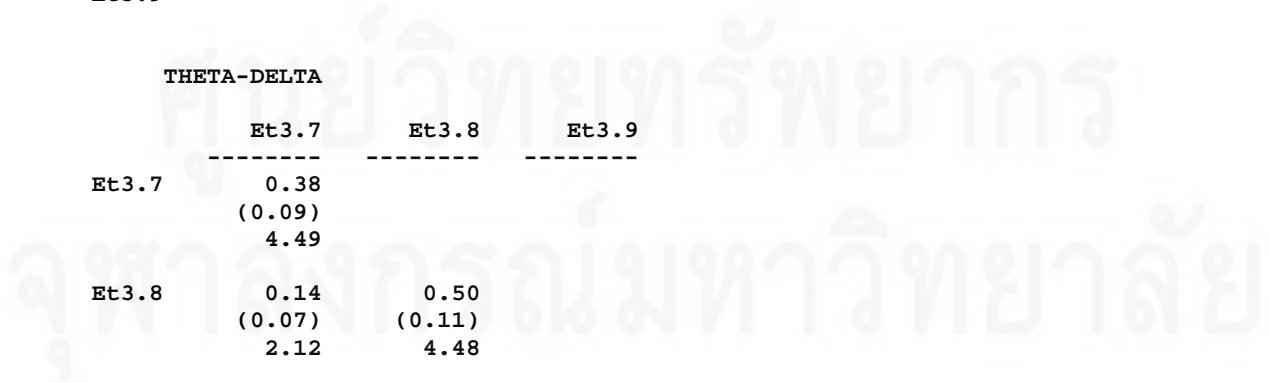
 1.00

THETA-DELTA

	Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6
Et3.1	0.23 (0.05) 4.40					
Et3.2	--	0.25 (0.07) 3.76				
Et3.3	--	--	0.39 (0.08) 4.72			
Et3.4	--	--	--	0.50 (0.11) 4.70		
Et3.5	--	--	--	0.13 (0.07) 1.81	0.37 (0.08) 4.47	
Et3.6	--	--	--	--	--	0.22 (0.06) 3.61
Et3.7	--	0.12 (0.06) 2.19	--	--	--	--
Et3.8	--	--	--	--	--	--
Et3.9	--	--	--	--	--	--

THETA-DELTA

	Et3.7	Et3.8	Et3.9
Et3.7	0.38 (0.09) 4.49		
Et3.8	0.14 (0.07) 2.12	0.50 (0.11) 4.48	
Et3.9	--	--	0.39 (0.09) 4.24



Squared Multiple Correlations for X - Variables

Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6
-----	-----	-----	-----	-----	-----
0.43	0.62	0.24	0.25	0.39	0.64

Squared Multiple Correlations for X - Variables

Et3.7	Et3.8	Et3.9
-----	-----	-----
0.40	0.39	0.49

Goodness of Fit Statistics

Degrees of Freedom = 24

Minimum Fit Function Chi-Square = 25.39 (P = 0.38)

Normal Theory Weighted Least Squares Chi-Square = 26.39 (P = 0.33)

Estimated Non-centrality Parameter (NCP) = 2.39

90 Percent Confidence Interval for NCP = (0.0 ; 19.14)

Minimum Fit Function Value = 0.52

Population Discrepancy Function Value (F0) = 0.049

90 Percent Confidence Interval for F0 = (0.0 ; 0.39)

Root Mean Square Error of Approximation (RMSEA) = 0.045

90 Percent Confidence Interval for RMSEA = (0.0 ; 0.13)

P-Value for Test of Close Fit (RMSEA < 0.05) = 0.49

Expected Cross-Validation Index (ECVI) = 1.40

90 Percent Confidence Interval for ECVI = (1.35 ; 1.74)

ECVI for Saturated Model = 1.84

ECVI for Independence Model = 7.60

Chi-Square for Independence Model with 36 Degrees of Freedom = 354.60

Independence AIC = 372.60

Model AIC = 68.39

Saturated AIC = 90.00

Independence CAIC = 398.81

Model CAIC = 129.54

Saturated CAIC = 221.04

Normed Fit Index (NFI) = 0.93

Non-Normed Fit Index (NNFI) = 0.99

Parsimony Normed Fit Index (PNFI) = 0.62

Comparative Fit Index (CFI) = 1.00

Incremental Fit Index (IFI) = 1.00

Relative Fit Index (RFI) = 0.89

Critical N (CN) = 83.95

Root Mean Square Residual (RMR) = 0.039

Standardized RMR = 0.061

Goodness of Fit Index (GFI) = 0.89

Adjusted Goodness of Fit Index (AGFI) = 0.80

Parsimony Goodness of Fit Index (PGFI) = 0.48

! MEASUREMENT OF ETHICALITY MODEL

Modification Indices and Expected Change

No Non-Zero Modification Indices for LAMBDA-X

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6
-------	-------	-------	-------	-------	-------

Et3.1	- -					
Et3.2	0.11	- -				
Et3.3	1.49	0.20	- -			
Et3.4	2.93	0.97	1.23	- -		
Et3.5	3.61	1.84	2.26	- -	- -	
Et3.6	0.27	2.09	0.19	0.59	0.88	- -
Et3.7	0.65	- -	0.61	0.19	0.01	0.07
Et3.8	2.72	2.73	1.48	3.25	0.69	0.40
Et3.9	0.10	0.19	0.01	0.43	1.57	0.32

Modification Indices for THETA-DELTA

	Et3.7	Et3.8	Et3.9
Et3.7	- -		
Et3.8	- -	- -	
Et3.9	1.75	2.70	- -

Expected Change for THETA-DELTA

	Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6
Et3.1	- -					
Et3.2	-0.01	- -				
Et3.3	0.06	-0.02	- -			
Et3.4	0.09	-0.05	0.07	- -		
Et3.5	-0.08	0.06	-0.08	- -	- -	
Et3.6	0.02	-0.07	0.02	0.04	0.05	- -
Et3.7	0.03	- -	0.04	0.02	0.01	-0.01
Et3.8	-0.09	0.11	-0.08	-0.12	0.05	-0.04
Et3.9	0.02	0.02	-0.01	-0.04	-0.07	0.03

Expected Change for THETA-DELTA

	Et3.7	Et3.8	Et3.9
Et3.7	- -		
Et3.8	- -	- -	
Et3.9	-0.07	0.11	- -

Maximum Modification Index is 3.61 for Element (5, 1) of THETA-DELTA

! MEASUREMENT OF ETHICALITY MODEL

Factor Scores Regressions

KSI

	Et3.1	Et3.2	Et3.3	Et3.4	Et3.5	Et3.6
ETHICALI	0.23	0.31	0.11	0.07	0.15	0.36

KSI

	Et3.7	Et3.8	Et3.9
ETHICALI	0.02	0.14	0.20

Time used: 0.047 Seconds

! MEASUREMENT OF META-EVALUATION MODEL

DA NI=5 NO=50
 LA FI=LA.t.DAT
 KM FI=KM.t.DAT
 SD FI=SD.t.DAT
 MO NX=5 NK=1 LX=FR TD=SY

FR TD(5,2) TD(4,3)
 LK
 META-EVALUATION
 PD
 OU MI FS

! MEASUREMENT OF META-EVALUATION MODEL

Number of Input Variables	5
Number of Y - Variables	0
Number of X - Variables	5
Number of ETA - Variables	0
Number of KSI - Variables	1
Number of Observations	50

! MEASUREMENT OF META-EVALUATION MODEL

Covariance Matrix

	Validity	Utility	Ethicali	Credibil	Cost-eff
	-----	-----	-----	-----	-----
Validity	1.00				
Utility	0.77	1.00			
Ethicali	0.38	0.35	1.00		
Credibil	0.70	0.74	0.43	1.00	
Cost-eff	0.15	0.30	0.14	0.13	1.00

! MEASUREMENT OF META-EVALUATION MODEL

Parameter Specifications

LAMBDA-X

	META-EVAL

Validity	1
Utility	2
Ethicali	3
Credibil	4
Cost-eff	5

THETA-DELTA

	Validity	Utility	Ethicali	Credibil	Cost-eff
	-----	-----	-----	-----	-----
Validity	6				
Utility	0	7			
Ethicali	0	0	8		
Credibil	0	0	9	10	
Cost-eff	0	11	0	0	12

! MEASUREMENT OF CREDIBILITY MODEL

DA NI=4 NO=50
 LA FI=LA04.DAT
 KM FI=KM04.DAT
 SD FI=SD04.DAT
 MO NX=4 NK=1 LX=FR TD=SY
 LK
 CREDIBILITY
 PD
 OU MI FS

! MEASUREMENT OF CREDIBILITY MODEL

Number of Input Variables	4
Number of Y - Variables	0

Number of X - Variables 4
 Number of ETA - Variables 0
 Number of KSI - Variables 1
 Number of Observations 50

! MEASUREMENT OF CREDIBILITY MODEL

Covariance Matrix

	Cr4.1	Cr4.2	Cr4.3	Cr4.4
Cr4.1	0.49			
Cr4.2	0.20	0.51		
Cr4.3	0.19	0.17	0.50	
Cr4.4	0.14	0.06	0.15	0.69

! MEASUREMENT OF CREDIBILITY MODEL

Parameter Specifications

LAMBDA-X

	CREDIBIL
Cr4.1	1
Cr4.2	2
Cr4.3	3
Cr4.4	4

THETA-DELTA

Cr4.1	Cr4.2	Cr4.3	Cr4.4
5	6	7	8

! MEASUREMENT OF CREDIBILITY MODEL

Number of Iterations = 5

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	CREDIBIL
Cr4.1	0.48 (0.12) 3.85
Cr4.2	0.39 (0.12) 3.26
Cr4.3	0.43 (0.12) 3.48
Cr4.4	0.28 (0.14) 1.96

PHI

CREDIBIL

1.00

THETA-DELTA

Cr4.1	Cr4.2	Cr4.3	Cr4.4
-----	-----	-----	-----
0.26	0.35	0.32	0.62
(0.10)	(0.09)	(0.09)	(0.13)
2.60	3.77	3.40	4.63

Squared Multiple Correlations for X - Variables

Cr4.1	Cr4.2	Cr4.3	Cr4.4
-----	-----	-----	-----
0.47	0.31	0.36	0.11

Goodness of Fit Statistics

Degrees of Freedom = 2
 Minimum Fit Function Chi-Square = 1.14 (P = 0.57)
 Normal Theory Weighted Least Squares Chi-Square = 1.10 (P = 0.58)
 Estimated Non-centrality Parameter (NCP) = 0.0
 90 Percent Confidence Interval for NCP = (0.0 ; 5.52)

Minimum Fit Function Value = 0.023
 Population Discrepancy Function Value (F0) = 0.0
 90 Percent Confidence Interval for F0 = (0.0 ; 0.11)
 Root Mean Square Error of Approximation (RMSEA) = 0.0
 90 Percent Confidence Interval for RMSEA = (0.0 ; 0.24)
 P-Value for Test of Close Fit (RMSEA < 0.05) = 0.62

Expected Cross-Validation Index (ECVI) = 0.37
 90 Percent Confidence Interval for ECVI = (0.37 ; 0.48)
 ECVI for Saturated Model = 0.41
 ECVI for Independence Model = 0.73

Chi-Square for Independence Model with 6 Degrees of Freedom = 27.65

Independence AIC = 35.65
 Model AIC = 17.10
 Saturated AIC = 20.00
 Independence CAIC = 47.30
 Model CAIC = 40.39
 Saturated CAIC = 49.12

Normed Fit Index (NFI) = 0.96
 Non-Normed Fit Index (NNFI) = 1.12
 Parsimony Normed Fit Index (PNFI) = 0.32
 Comparative Fit Index (CFI) = 1.00
 Incremental Fit Index (IFI) = 1.03
 Relative Fit Index (RFI) = 0.88

Critical N (CN) = 397.26

Root Mean Square Residual (RMR) = 0.021
 Standardized RMR = 0.036
 Goodness of Fit Index (GFI) = 0.99
 Adjusted Goodness of Fit Index (AGFI) = 0.94
 Parsimony Goodness of Fit Index (PGFI) = 0.20

! MEASUREMENT OF CREDIBILITY MODEL

Modification Indices and Expected Change

No Non-Zero Modification Indices for LAMBDA-X

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

	Cr4.1	Cr4.2	Cr4.3	Cr4.4
Cr4.1	- -			
Cr4.2	0.47	- -		
Cr4.3	1.01	0.06	- -	
Cr4.4	0.06	1.01	0.47	- -

Expected Change for THETA-DELTA

	Cr4.1	Cr4.2	Cr4.3	Cr4.4
Cr4.1	- -			
Cr4.2	0.09	- -		
Cr4.3	-0.15	0.03	- -	
Cr4.4	0.02	-0.08	0.06	- -

Maximum Modification Index is 1.01 for Element (3, 1) of THETA-DELTA

! MEASUREMENT OF CREDIBILITY MODEL

Factor Scores Regressions

KSI

	Cr4.1	Cr4.2	Cr4.3	Cr4.4
CREDIBIL	0.61	0.37	0.44	0.15

Time used: 0.063 Seconds

! MEASUREMENT OF COST-EFFECTIVENESS MODEL

DA NI=2 NO=50
 LA FI=LA05.DAT
 KM FI=KM05.DAT
 SD FI=SD05.DAT
 MO NX=2 NK=1 LX=FR TD=SY
 FI TD(1,1) TD(2,2)
 VA 0.25 TD(2,2)
 LK
 COST-EFFECTIVENESS
 PD
 OU MI FS

! MEASUREMENT OF COST-EFFECTIVENESS MODEL

Number of Input Variables 2
 Number of Y - Variables 0
 Number of X - Variables 2
 Number of ETA - Variables 0
 Number of KSI - Variables 1
 Number of Observations 50

! MEASUREMENT OF COST-EFFECTIVENESS MODEL

Covariance Matrix

	Ce5.1	Ce5.2
Ce5.1	0.59	
Ce5.2	0.19	0.31

! MEASUREMENT OF COST-EFFECTIVENESS MODEL

Parameter Specifications

```

LAMBDA-X
      COST-EFF
      -----
Ce5.1      1
Ce5.2      2
    
```

! MEASUREMENT OF COST-EFFECTIVENESS MODEL

Number of Iterations = 1

LISREL Estimates (Maximum Likelihood)

```

LAMBDA-X
      COST-EFF
      -----
Ce5.1      0.77
            (0.08)
            9.90
Ce5.2      0.24
            (0.08)
            3.20
    
```

PHI

```

      COST-EFF
      -----
            1.00
    
```

THETA-DELTA

```

      Ce5.1      Ce5.2
      -----      -----
      - -      0.25
    
```

Squared Multiple Correlations for X - Variables

```

      Ce5.1      Ce5.2
      -----      -----
            1.00      0.19
    
```

Goodness of Fit Statistics

Degrees of Freedom = 1
 Minimum Fit Function Chi-Square = 0.0013 (P = 0.97)
 Normal Theory Weighted Least Squares Chi-Square = 0.0013 (P = 0.97)

The Fit is Perfect !

! MEASUREMENT OF COST-EFFECTIVENESS MODEL

Modification Indices and Expected Change

No Non-Zero Modification Indices for LAMBDA-X

No Non-Zero Modification Indices for PHI

No Non-Zero Modification Indices for THETA-DELTA

Maximum Modification Index is 0.00 for Element (2, 2) of THETA-DELTA

! MEASUREMENT OF COST-EFFECTIVENESS MODEL

Factor Scores Regressions

KSI

	Ce5.1	Ce5.2
COST-EFF	1.30	- -

Time used: 0.063 Seconds

! MEASUREMENT OF META-EVALUATION MODEL

Number of Iterations = 7

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

	META-EVAL
Validity	0.86 (0.12) 7.16
Utility	0.90 (0.12) 7.68
Ethicali	0.40 (0.15) 2.75
Credibil	0.82 (0.12) 6.80
Cost-eff	0.17 (0.16) 1.07

PHI

META-EVAL
1.00

THETA-DELTA

	Validity	Utility	Ethicali	Credibil	Cost-eff
Validity	0.27 (0.08) 3.25				
Utility	- -	0.19 (0.08)			

			2.40		
Ethicali	- -	- -	0.84 (0.17) 4.82		
Credibil	- -	- -	0.10 (0.09) 1.17	0.32 (0.09) 3.68	
Cost-eff	- -	0.15 (0.09) 1.71	- -	- -	0.97 (0.20) 4.92

Squared Multiple Correlations for X - Variables

Validity	Utility	Ethicali	Credibil	Cost-eff
-----	-----	-----	-----	-----
0.73	0.81	0.16	0.68	0.03

Goodness of Fit Statistics

Degrees of Freedom = 3
 Minimum Fit Function Chi-Square = 0.84 (P = 0.84)
 Normal Theory Weighted Least Squares Chi-Square = 0.84 (P = 0.84)
 Estimated Non-centrality Parameter (NCP) = 0.0
 90 Percent Confidence Interval for NCP = (0.0 ; 2.74)

Minimum Fit Function Value = 0.017
 Population Discrepancy Function Value (F0) = 0.0
 90 Percent Confidence Interval for F0 = (0.0 ; 0.056)
 Root Mean Square Error of Approximation (RMSEA) = 0.0
 90 Percent Confidence Interval for RMSEA = (0.0 ; 0.14)
 P-Value for Test of Close Fit (RMSEA < 0.05) = 0.86

Expected Cross-Validation Index (ECVI) = 0.55
 90 Percent Confidence Interval for ECVI = (0.55 ; 0.61)
 ECVI for Saturated Model = 0.61
 ECVI for Independence Model = 2.44

Chi-Square for Independence Model with 10 Degrees of Freedom = 109.54
 Independence AIC = 119.54
 Model AIC = 24.84
 Saturated AIC = 30.00
 Independence CAIC = 134.11
 Model CAIC = 59.78
 Saturated CAIC = 73.68

Normed Fit Index (NFI) = 0.99
 Non-Normed Fit Index (NNFI) = 1.07
 Parsimony Normed Fit Index (PNFI) = 0.30
 Comparative Fit Index (CFI) = 1.00
 Incremental Fit Index (IFI) = 1.02
 Relative Fit Index (RFI) = 0.97

Critical N (CN) = 664.96

Root Mean Square Residual (RMR) = 0.021
 Standardized RMR = 0.021
 Goodness of Fit Index (GFI) = 0.99
 Adjusted Goodness of Fit Index (AGFI) = 0.97
 Parsimony Goodness of Fit Index (PGFI) = 0.20

! MEASUREMENT OF META-EVALUATION MODEL

Modification Indices and Expected Change

No Non-Zero Modification Indices for LAMBDA-X

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

	Validity	Utility	Ethicali	Credibil	Cost-eff
	-----	-----	-----	-----	-----
Validity	- -				
Utility	0.01	- -			
Ethicali	0.41	0.52	- -		
Credibil	0.41	0.52	- -	- -	
Cost-eff	0.01	- -	0.54	0.15	- -

Expected Change for THETA-DELTA

	Validity	Utility	Ethicali	Credibil	Cost-eff
	-----	-----	-----	-----	-----
Validity	- -				
Utility	-0.06	- -			
Ethicali	0.05	-0.06	- -		
Credibil	-0.11	0.12	- -	- -	
Cost-eff	0.01	- -	0.09	-0.04	- -

Maximum Modification Index is 0.54 for Element (5, 3) of THETA-DELTA

! MEASUREMENT OF METAEVALUATION MODEL

Factor Scores Regressions

KSI

	Validity	Utility	Ethicali	Credibil	Cost-eff
	-----	-----	-----	-----	-----
METAEVAL	0.31	0.50	0.02	0.24	-0.06

Time used: 0.047 Seconds

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Print out การวิเคราะห์ G-Coefficient

```

GGGGGGGGG EEEEEEEEEEE NN NN 00000000 VV VV AAAAAAAAAA
GGGGGGGGGG EEEEEEEEEEE NNN NN 0000000000 VV VV AAAAAAAAAA
GG G EEE EEE NNN NN 00 00 VV VV AA AA
GG EE NN NN NN 00 00 VV VV AA AA
GG EE NN NN NN 00 00 VV VV AA AA
GG EE NN NN NN 00 00 VV VV AA AA
GG EE NN NN NN 00 00 VV VV AA AA
GG EEEEEEEE NN NN NN 00 00 VV VV AA AA
GG GGGG EEEEEEEE NN NNN 00 00 VV VV AAAAAAAAAA
GG GGGG EE NN NNN 00 00 VV VV AAAAAAAAAA
GG GG EE NN NN 00 00 VV VV AA AA
GG GG EE NN NN 00 00 VV VV AA AA
GG GG EE NN NN 00 00 V V V AA AA
GG GG EE NN NN 00 00 VVVV AA AA
GGGGGGGGGG EEEEEEEEEEE NN NN 0000000000 VVVV AA AA
GGGGGGGGGG EEEEEEEEEEE NN NN 00000000 VV AA AA

```

A GENERAL PURPOSE ANALYSIS OF VARIANCE SYSTEM

GENOVA IS A FORTRAN 77 PROGRAM FOR ANALYSIS OF VARIANCE AND GENERALIZABILITY ANALYSES WITH BALANCED DESIGNS

AUTHORS

Joe E. Crick, Ed.D.
 Chief Technology & Information Officer
 Vice President Applications and Database Services
 National Board of Medical Examiners
 Philadelphia, PA 19104

Robert L. Brennan, Ed.D.
 Director, Iowa Testing Program
 University of Iowa
 Iowa City, Iowa 52242

VERSION 3.1
January, 2001

GENOVA has been checked for accuracy of output, however the authors can make no assurances that the program is totally without error.

GENOVA was developed in part under contract No. N00123-78-C-1206 with the Navy Personnel Research and Development Center (NPRDC); Robert L. Brennan Principal Investigator. GENOVA does not necessarily reflect NPRDC positions or policy, and no official endorsement should be inferred

GENOVA VERSION 3.1 PAGE 1

CONTROL CARD INPUT LISTING

```

COLUMN 1111111112222222223333333334444444445555555556666666667777777778
1234567890123456789012345678901234567890123456789012345678901234567890
STUDY DATA SET 1 - RANDOM MODEL - I X (R:P) DESIGN - RUN 1
COMMENT # RECORDS = 38
COMMENT # VALUES PER RECORD = 10
COMMENT
OPTIONS RECORDS ALL CORRELATION
EFFECT * I 38 0
EFFECT + P 5 0
EFFECT + R:P 2 0
FORMAT (10F2.0)
PROCESS

```

GENOVA VERSION 3.1 PAGE 2

DATA SET 1 - RANDOM MODEL - I X (R:P) DESIGN - RUN 1

EXPANDED MAIN AND INTERACTION EFFECT TABLE

(** = INFINITE)	I	P	R	PRIMARY	TOTAL	DEGREES
SAMPLE SIZE	38	5	2	INDICES	NUMBER	OF
UNIVERSE SIZE	****	****	****		INDICES	FREEDOM
* I	* 1	* 0	* 0	1	1	37
* P	* 0	* 1	* 0	1	1	4
* R:P	* 0	* 2	* 1	1	2	5

* IP	* 1	* 1	* 0	2	2	148
* IR:P	* 1	* 2	* 1	2	3	185

GENOVA VERSION 3.1 PAGE 3

DATA SET 1 - RANDOM MODEL - I X (R:P) DESIGN - RUN 1

INPUT RECORD LISTING WITH RECORD MEANS

RECORD #	1	2	3	4	5	6	7
RECORD # 1	4.0000 3.0000	4.0000 2.0000	1.0000 2.6000	2.0000	3.0000	3.0000	2.0000
RECORD # 2	4.0000 3.0000	4.0000 2.0000	2.0000 2.5000	3.0000	1.0000	3.0000	1.0000
RECORD # 3	2.0000 3.0000	3.0000 1.0000	3.0000 2.4000	4.0000	4.0000	2.0000	1.0000
RECORD # 4	2.0000 4.0000	2.0000 3.0000	2.0000 2.8000	4.0000	4.0000	3.0000	2.0000
RECORD # 5	3.0000 3.0000	2.0000 3.0000	3.0000 2.5000	3.0000	3.0000	2.0000	1.0000
RECORD # 6	4.0000 2.0000	2.0000 3.0000	2.0000 2.6000	3.0000	4.0000	2.0000	2.0000
RECORD # 7	4.0000	3.0000	1.0000	3.0000	3.0000	3.0000	2.0000

		3. 00000	2. 00000	2. 70000					
RECORD #	8	2. 00000 2. 00000	2. 00000 1. 00000	1. 00000 1. 70000	2. 00000	2. 00000	2. 00000	1. 00000	2. 00000
RECORD #	9	3. 00000 4. 00000	3. 00000 1. 00000	4. 00000 2. 70000	4. 00000	3. 00000	1. 00000	2. 00000	2. 00000
RECORD #	10	3. 00000 3. 00000	3. 00000 2. 00000	3. 00000 2. 60000	3. 00000	4. 00000	1. 00000	2. 00000	2. 00000
RECORD #	11	3. 00000 3. 00000	3. 00000 3. 00000	2. 00000 2. 70000	3. 00000	3. 00000	3. 00000	1. 00000	3. 00000
RECORD #	12	3. 00000 3. 00000	1. 00000 3. 00000	4. 00000 2. 30000	1. 00000	1. 00000	3. 00000	2. 00000	2. 00000
RECORD #	13	4. 00000 3. 00000	4. 00000 4. 00000	4. 00000 3. 60000	4. 00000	3. 00000	3. 00000	4. 00000	3. 00000
RECORD #	14	4. 00000 3. 00000	3. 00000 2. 00000	3. 00000 2. 40000	3. 00000	1. 00000	3. 00000	1. 00000	1. 00000
RECORD #	15	4. 00000 3. 00000	4. 00000 4. 00000	4. 00000 3. 70000	4. 00000	4. 00000	3. 00000	3. 00000	4. 00000
RECORD #	16	4. 00000 2. 00000	2. 00000 3. 00000	4. 00000 2. 80000	4. 00000	2. 00000	2. 00000	3. 00000	2. 00000
RECORD #	17	4. 00000 3. 00000	4. 00000 2. 00000	4. 00000 3. 20000	4. 00000	2. 00000	3. 00000	3. 00000	3. 00000

GENOVA VERSION 3.1
G STUDY

DATA SET 1 - RANDOM MODEL - I X (R:P) DESIGN - RUN 1

PAGE 4

INPUT RECORD LISTING WITH RECORD MEANS

RECORD #	18	1. 00000 3. 00000	2. 00000 1. 00000	1. 00000 1. 90000	1. 00000	1. 00000	4. 00000	3. 00000	2. 00000
RECORD #	19	1. 00000 2. 00000	2. 00000 1. 00000	1. 00000 1. 50000	1. 00000	1. 00000	1. 00000	2. 00000	3. 00000
RECORD #	20	3. 00000 3. 00000	4. 00000 1. 00000	4. 00000 2. 70000	4. 00000	3. 00000	1. 00000	2. 00000	2. 00000
RECORD #	21	3. 00000 3. 00000	4. 00000 2. 00000	3. 00000 2. 80000	4. 00000	4. 00000	2. 00000	1. 00000	2. 00000
RECORD #	22	3. 00000 4. 00000	4. 00000 3. 00000	3. 00000 3. 40000	4. 00000	4. 00000	3. 00000	3. 00000	3. 00000
RECORD #	23	3. 00000 4. 00000	3. 00000 4. 00000	4. 00000 3. 10000	1. 00000	2. 00000	4. 00000	2. 00000	4. 00000
RECORD #	24	4. 00000 3. 00000	4. 00000 3. 00000	4. 00000 3. 70000	4. 00000	4. 00000	4. 00000	4. 00000	3. 00000
RECORD #	25	4. 00000 4. 00000	4. 00000 4. 00000	4. 00000 3. 40000	4. 00000	4. 00000	1. 00000	2. 00000	3. 00000
RECORD #	26	1. 00000 1. 00000	2. 00000 1. 00000	1. 00000 1. 70000	1. 00000	2. 00000	1. 00000	4. 00000	3. 00000
RECORD #	27	2. 00000 4. 00000	4. 00000 1. 00000	1. 00000 2. 70000	4. 00000	1. 00000	3. 00000	4. 00000	3. 00000
RECORD #	28	1. 00000 2. 00000	2. 00000 2. 00000	2. 00000 2. 20000	2. 00000	4. 00000	2. 00000	3. 00000	2. 00000
RECORD #	29	4. 00000 4. 00000	4. 00000 4. 00000	4. 00000 3. 70000	4. 00000	2. 00000	4. 00000	4. 00000	3. 00000
RECORD #	30	4. 00000 1. 00000	1. 00000 1. 00000	1. 00000 1. 60000	1. 00000	4. 00000	1. 00000	1. 00000	1. 00000
RECORD #	31	2. 00000 4. 00000	4. 00000 3. 00000	4. 00000 3. 20000	4. 00000	3. 00000	3. 00000	2. 00000	3. 00000
RECORD #	32	3. 00000 2. 00000	4. 00000 4. 00000	4. 00000 3. 10000	4. 00000	3. 00000	2. 00000	2. 00000	3. 00000
RECORD #	33	3. 00000 4. 00000	4. 00000 3. 00000	4. 00000 3. 60000	3. 00000	4. 00000	4. 00000	3. 00000	4. 00000
RECORD #	34	2. 00000 2. 00000	3. 00000 4. 00000	4. 00000 3. 10000	4. 00000	3. 00000	3. 00000	4. 00000	2. 00000

GENOVA VERSION 3.1
G STUDY

DATA SET 1 - RANDOM MODEL - I X (R:P) DESIGN - RUN 1

PAGE 5

INPUT RECORD LISTING WITH RECORD MEANS

RECORD #	35	4. 00000 3. 00000	2. 00000 3. 00000	3. 00000 2. 70000	3. 00000	2. 00000	2. 00000	2. 00000	3. 00000
RECORD #	36	2. 00000 2. 00000	4. 00000 2. 00000	4. 00000 3. 00000	4. 00000	4. 00000	1. 00000	4. 00000	3. 00000
RECORD #	37	3. 00000 1. 00000	4. 00000 3. 00000	2. 00000 2. 30000	2. 00000	2. 00000	2. 00000	2. 00000	2. 00000
RECORD #	38	2. 00000 1. 00000	1. 00000 1. 00000	1. 00000 1. 30000	1. 00000	2. 00000	1. 00000	2. 00000	1. 00000

*** GRAND MEAN = 2.6973684 ***

MEAN SCORES FOR EFFECT: P SUBSCRIPT NOTATION: (P)
(1) = 2.986842 (2) = 2.894737 (3) = 2.592105 (4) = 2.394737
(5) = 2.618421

MEAN SCORES FOR EFFECT: R: P SUBSCRIPT NOTATION: (P, R)
(1, 1) = 2.947368 (1, 2) = 3.026316
(2, 1) = 2.789474 (2, 2) = 3.000000
(3, 1) = 2.789474 (3, 2) = 2.394737
(4, 1) = 2.342105 (4, 2) = 2.447368
(5, 1) = 2.815789 (5, 2) = 2.421053

(** = INFINITE) I P R
SAMPLE SIZE 38 5 2
UNIVERSE SIZE ****

EFFECT	DEGREES OF FREEDOM	SUMS OF SQUARES FOR MEAN SCORES	SUMS OF SQUARES FOR SCORE EFFECTS	MEAN SQUARES	F STATISTIC	(QF = QUASI F RATIO) F-TEST DEGREES OF FREEDOM NUMERATOR	OF FREEDOM DENOMINATOR
I	37	2912.90000	148.09737	4.00263	5.15145	37	148
P	4	2782.40789	17.60526	4.40132	3.04113 QF	4 QF	5 QF
R: P	5	2789.50000	7.09211	1.41842	1.89590	5	185
IP	148	3045.50000	114.99474	0.77699	1.03855	148	185
IR: P	185	3191.00000	138.40789	0.74815			
MEAN		2764.80263					
TOTAL	379		426.19737				

NOTE: FOR GENERALIZABILITY ANALYSES, F-STATISTICS SHOULD BE IGNORED

(** = INFINITE) I P R
SAMPLE SIZE 38 5 2
UNIVERSE SIZE ****

QFM = QUADRATIC FORM

EFFECT	DEGREES OF FREEDOM	MODEL USING ALGORITHM	VARIANCE USING EMS EQUATIONS	COMPONENTS STANDARD ERROR
I	37	0.3225640	0.3225640	0.0910847
P	4	0.0388691	0.0388691	0.0349269
R: P	5	0.0176387	0.0176387	0.0200557
IP	148	0.0144203	0.0144203	0.0592367
IR: P	185	0.7481508	0.7481508	0.0773719

NOTE: THE "ALGORITHM" AND "EMS" ESTIMATED VARIANCE COMPONENTS WILL BE IDENTICAL IF THERE ARE NO NEGATIVE ESTIMATES

(** = INFINITE) I P R
SAMPLE SIZE 38 5 2
UNIVERSE SIZE ****

EMS(I) = 1.00*VC(IR: P) + 2.00*VC(IP) + 10.00*VC(I)

EMS(P) = 1.00*VC(IR: P) + 2.00*VC(IP) + 38.00*VC(R: P) + 76.00*VC(P)

EMS(R: P) = 1.00*VC(IR: P) + 38.00*VC(R: P)

EMS(IP) = 1.00*VC(IR: P) + 2.00*VC(IP)

EMS(IR: P) = 1.00*VC(IR: P)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000106	0.0012199			
R: P	0.0000000	-0.0002011	0.0004022		
IP	-0.0004025	-0.0000923	0.0000788	0.0035090	
IR: P	0.0000000	0.0000788	-0.0001575	-0.0029932	0.0059864

11

G STUDY

DATA SET 1 - RANDOM MODEL - I X (R: P) DESIGN - RUN 1

CORRELATION MATRIX FOR ESTIMATED VARIANCE COMPONENTS

	I	P	R: P	IP	IR: P
I	1.0000000				
P	0.0033293	1.0000000			
R: P	0.0000000	-0.2871095	1.0000000		
IP	-0.0745942	-0.0446321	0.0663017	1.0000000	
IR: P	0.0000000	0.0291481	-0.1015225	-0.6530739	1.0000000

CONTROL CARD INPUT LISTING

12

COLUMN 11111111112222222222333333333344444444445555555555666666666677777777778
12345678901234567890123456789012345678901234567890123456789012345678901234567890
DSTUDY FIRST SET OF D STUDY CONTROL CARDS
DEFFECT \$ I 38
DEFFECT P 3 5 7 9 3 5 7 9 3 5 7 9
DEFFECT R: P 1 1 1 2 2 2 2 3 3 3 3
ENDDSTUDY

D STUDY

FIRST SET OF D STUDY CONTROL CARDS

13

D STUDY DESIGN NUMBER 001-001

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 3 1

EFFECT	VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES					VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES				
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	VARIANCE COMPONENTS FOR MEAN SCORES ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	VARIANCE COMPONENTS FOR MEAN SCORES ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	3	0.01296	0.01164	0.03887	1.0000	3	0.01296	0.01164
R: P	0.01764	1.0000	3	0.00588	0.00669	0.01764	1.0000	3	0.00588	0.00669
IP	0.01442	1.0000	3	0.00481	0.01975	0.01442	1.0000	3	0.00481	0.01975
IR: P	0.74815	1.0000	3	0.24938	0.02579	0.74815	1.0000	3	0.24938	0.02579

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.57675	0.75944	0.09175	
LOWER CASE DELTA	0.25419	0.50417	0.01975	GENERALIZABILITY COEFFICIENT = 0.55927 (1.26899)
UPPER CASE DELTA	0.27303	0.52252	0.02247	PHI = 0.54159 (1.18144)
MEAN	0.03401	0.18443		

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

D STUDY

FIRST SET OF D STUDY CONTROL CARDS

14

D STUDY DESIGN NUMBER 001-001

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000035	0.0001355			
R: P	0.0000000	-0.0000223	0.0000447		
IP	-0.0001342	-0.0000103	0.0000088	0.0003899	
IR: P	0.0000000	0.0000088	-0.0000175	-0.0003326	0.0006652

D STUDY

FIRST SET OF D STUDY CONTROL CARDS

15

D STUDY DESIGN NUMBER 001-002

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 5 1

EFFECT	VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES					VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES				
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	VARIANCE COMPONENTS FOR MEAN SCORES ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	VARIANCE COMPONENTS FOR MEAN SCORES ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	5	0.00777	0.00699	0.03887	1.0000	5	0.00777	0.00699
R: P	0.01764	1.0000	5	0.00353	0.00401	0.01764	1.0000	5	0.00353	0.00401
IP	0.01442	1.0000	5	0.00288	0.01185	0.01442	1.0000	5	0.00288	0.01185
IR: P	0.74815	1.0000	5	0.14963	0.01547	0.74815	1.0000	5	0.14963	0.01547

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.47508	0.68926	0.09097	
LOWER CASE DELTA	0.15251	0.39053	0.01185	
UPPER CASE DELTA	0.16382	0.40474	0.01348	
MEAN	0.02380	0.15428		

GENERALIZABILITY COEFFICIENT = 0.67897 (2.11498)
PHI = 0.66319 (1.96907)

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 16
D STUDY DESIGN NUMBER 001-002

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000021	0.0000488			
R: P	0.0000000	-0.0000080	0.0000161		
IP	-0.0000805	-0.0000037	0.0000032	0.0001404	
IR: P	0.0000000	0.0000032	-0.0000063	-0.0001197	0.0002395

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 17
D STUDY DESIGN NUMBER 001-003

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 7 1

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES			VARIANCE COMPONENTS FOR MEAN SCORES						
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	7	0.00555	0.00499	0.03887	1.0000	7	0.00555	0.00499
R: P	0.01764	1.0000	7	0.00252	0.00287	0.01764	1.0000	7	0.00252	0.00287
IP	0.01442	1.0000	7	0.00206	0.00846	0.01442	1.0000	7	0.00206	0.00846
IR: P	0.74815	1.0000	7	0.10688	0.01105	0.74815	1.0000	7	0.10688	0.01105

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.43150	0.65689	0.09085	
LOWER CASE DELTA	0.10894	0.33006	0.00846	
UPPER CASE DELTA	0.11701	0.34207	0.00963	
MEAN	0.01943	0.13938		

GENERALIZABILITY COEFFICIENT = 0.74754 (2.96097)
PHI = 0.73381 (2.75669)

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 18
D STUDY DESIGN NUMBER 001-003

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000015	0.0000249			
R: P	0.0000000	-0.0000041	0.0000082		
IP	-0.0000575	-0.0000019	0.0000016	0.0000716	
IR: P	0.0000000	0.0000016	-0.0000032	-0.0000611	0.0001222

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 19
D STUDY DESIGN NUMBER 001-004

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 9 1

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES			VARIANCE COMPONENTS FOR MEAN SCORES						
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	9	0.00432	0.00388	0.03887	1.0000	9	0.00432	0.00388
R: P	0.01764	1.0000	9	0.00196	0.00223	0.01764	1.0000	9	0.00196	0.00223
IP	0.01442	1.0000	9	0.00160	0.00658	0.01442	1.0000	9	0.00160	0.00658
IR: P	0.74815	1.0000	9	0.08313	0.00860	0.74815	1.0000	9	0.08313	0.00860

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.40729	0.63820	0.09083	
LOWER CASE DELTA	0.08473	0.29108	0.00658	GENERALIZABILITY COEFFICIENT = 0.79197 (3.80696)
UPPER CASE DELTA	0.09101	0.30168	0.00749	PHI = 0.77994 (3.54432)
MEAN	0.01700	0.13037		

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 20
D STUDY DESIGN NUMBER 001-004

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R:P	IP	IR:P
I	0.0082964				
P	0.0000012	0.0000151			
R:P	0.0000000	-0.0000025	0.0000050		
IP	-0.0000447	-0.0000011	0.0000010	0.0000433	
IR:P	0.0000000	0.0000010	-0.0000019	-0.0000370	0.0000739

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 21
D STUDY DESIGN NUMBER 001-005

OBJECT OF MEASUREMENT : I FACETS : P R:P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 3 2

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES			VARIANCE COMPONENTS FOR MEAN SCORES						
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	3	0.01296	0.01164	0.03887	1.0000	3	0.01296	0.01164
R:P	0.01764	1.0000	6	0.00294	0.00334	0.01764	1.0000	6	0.00294	0.00334
IP	0.01442	1.0000	3	0.00481	0.01975	0.01442	1.0000	3	0.00481	0.01975
IR:P	0.74815	1.0000	6	0.12469	0.01290	0.74815	1.0000	6	0.12469	0.01290

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.45206	0.67236	0.09084	
LOWER CASE DELTA	0.12950	0.35986	0.01495	GENERALIZABILITY COEFFICIENT = 0.71354 (2.49087)
UPPER CASE DELTA	0.14539	0.38131	0.01834	PHI = 0.68930 (2.21854)
MEAN	0.02779	0.16671		

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 22
D STUDY DESIGN NUMBER 001-005

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R:P	IP	IR:P
I	0.0082964				
P	0.0000035	0.0001355			
R:P	0.0000000	-0.0000112	0.0000112		
IP	-0.0001342	-0.0000103	0.0000044	0.0003899	
IR:P	0.0000000	0.0000044	-0.0000044	-0.0001663	0.0001663

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 23
D STUDY DESIGN NUMBER 001-006

OBJECT OF MEASUREMENT : I FACETS : P R:P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 5 2

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES			VARIANCE COMPONENTS FOR MEAN SCORES						
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	5	0.00777	0.00699	0.03887	1.0000	5	0.00777	0.00699
R:P	0.01764	1.0000	10	0.00176	0.00201	0.01764	1.0000	10	0.00176	0.00201
IP	0.01442	1.0000	5	0.00288	0.01185	0.01442	1.0000	5	0.00288	0.01185
IR:P	0.74815	1.0000	10	0.07482	0.00774	0.74815	1.0000	10	0.07482	0.00774

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.40026	0.63266	0.09064	
LOWER CASE DELTA	0.07770	0.27875	0.00897	
UPPER CASE DELTA	0.08724	0.29536	0.01100	
MEAN	0.02007	0.14167		

GENERALIZABILITY COEFFICIENT = 0.80588 (4.15145)
 PHI = 0.78712 (3.69757)

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 24

D STUDY DESIGN NUMBER 001-006

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R:P	IP	IR:P
I	0.0082964				
P	0.0000021	0.0000488			
R:P	0.0000000	-0.0000040	0.0000040		
IP	-0.0000805	-0.0000037	0.0000016	0.0001404	
IR:P	0.0000000	0.0000016	-0.0000016	-0.0000599	0.0000599

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 25

D STUDY DESIGN NUMBER 001-007

OBJECT OF MEASUREMENT : I FACETS : P R:P
 G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
 D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
 D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 7 2

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS			VARIANCE COMPONENTS FOR MEAN SCORES		VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS			VARIANCE COMPONENTS FOR MEAN SCORES	
	VARIANCE	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	7	0.00555	0.00499	0.03887	1.0000	7	0.00555	0.00499
R:P	0.01764	1.0000	14	0.00126	0.00143	0.01764	1.0000	14	0.00126	0.00143
IP	0.01442	1.0000	7	0.00206	0.00846	0.01442	1.0000	7	0.00206	0.00846
IR:P	0.74815	1.0000	14	0.05344	0.00553	0.74815	1.0000	14	0.05344	0.00553

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.37806	0.61487	0.09068	
LOWER CASE DELTA	0.05550	0.23558	0.00641	
UPPER CASE DELTA	0.06231	0.24962	0.00786	
MEAN	0.01676	0.12947		

GENERALIZABILITY COEFFICIENT = 0.85320 (5.81203)
 PHI = 0.83810 (5.17659)

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 26

D STUDY DESIGN NUMBER 001-007

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R:P	IP	IR:P
I	0.0082964				
P	0.0000015	0.0000249			
R:P	0.0000000	-0.0000021	0.0000021		
IP	-0.0000575	-0.0000019	0.0000008	0.0000716	
IR:P	0.0000000	0.0000008	-0.0000008	-0.0000305	0.0000305

GENOVA VERSION 3.1 D STUDY FIRST SET OF D STUDY CONTROL CARDS PAGE 27

D STUDY DESIGN NUMBER 001-008

OBJECT OF MEASUREMENT : I FACETS : P R:P
 G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
 D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
 D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 9 2

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS			VARIANCE COMPONENTS FOR MEAN SCORES		VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS			VARIANCE COMPONENTS FOR MEAN SCORES	
	VARIANCE	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	9	0.00432	0.00388	0.03887	1.0000	9	0.00432	0.00388
R:P	0.01764	1.0000	18	0.00098	0.00111	0.01764	1.0000	18	0.00098	0.00111
IP	0.01442	1.0000	9	0.00160	0.00658	0.01442	1.0000	9	0.00160	0.00658
IR:P	0.74815	1.0000	18	0.04156	0.00430	0.74815	1.0000	18	0.04156	0.00430

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.36573	0.60476	0.09073	
LOWER CASE DELTA	0.04317	0.20776	0.00498	GENERALIZABILITY COEFFICIENT = 0.88197 (7.47261)
UPPER CASE DELTA	0.04846	0.22015	0.00611	PHI = 0.86938 (6.65562)
MEAN	0.01492	0.12216		

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 28
D STUDY

D STUDY DESIGN NUMBER 001-008

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000012	0.0000151			
R: P	0.0000000	-0.0000012	0.0000012		
IP	-0.0000447	-0.0000011	0.0000005	0.0000433	
IR: P	0.0000000	0.0000005	-0.0000005	-0.0000185	0.0000185

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 29
D STUDY

D STUDY DESIGN NUMBER 001-009

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 3 3

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES			VARIANCE COMPONENTS FOR MEAN SCORES						
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	3	0.01296	0.01164	0.03887	1.0000	3	0.01296	0.01164
R: P	0.01764	1.0000	9	0.00196	0.00223	0.01764	1.0000	9	0.00196	0.00223
IP	0.01442	1.0000	3	0.00481	0.01975	0.01442	1.0000	3	0.00481	0.01975
IR: P	0.74815	1.0000	9	0.08313	0.00860	0.74815	1.0000	9	0.08313	0.00860

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.41050	0.64070	0.09094	
LOWER CASE DELTA	0.08793	0.29654	0.01556	GENERALIZABILITY COEFFICIENT = 0.78579 (3.66822)
UPPER CASE DELTA	0.10285	0.32070	0.01884	PHI = 0.75823 (3.13623)
MEAN	0.02572	0.16037		

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 30
D STUDY

D STUDY DESIGN NUMBER 001-009

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000035	0.0001355			
R: P	0.0000000	-0.0000074	0.0000050		
IP	-0.0001342	-0.0000103	0.0000029	0.0003899	
IR: P	0.0000000	0.0000029	-0.0000019	-0.0001109	0.0000739

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 31
D STUDY

D STUDY DESIGN NUMBER 001-010

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 5 3

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES			VARIANCE COMPONENTS FOR MEAN SCORES						
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	5	0.00777	0.00699	0.03887	1.0000	5	0.00777	0.00699
R: P	0.01764	1.0000	15	0.00118	0.00134	0.01764	1.0000	15	0.00118	0.00134
IP	0.01442	1.0000	5	0.00288	0.01185	0.01442	1.0000	5	0.00288	0.01185
IR: P	0.74815	1.0000	15	0.04988	0.00516	0.74815	1.0000	15	0.04988	0.00516

QFM = QUADRATIC FORM

		VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108		
EXPECTED OBSERVED SCORE	0.37532	0.61264	0.09068		
LOWER CASE DELTA	0.05276	0.22970	0.00934		
UPPER CASE DELTA	0.06171	0.24842	0.01130		
MEAN	0.01883	0.13721			
				GENERALIZABILITY COEFFICIENT = 0.85943 (6.11371)	
				PHI = 0.83941 (5.22705)	

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 32
D STUDY

D STUDY DESIGN NUMBER 001-010

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000021	0.0000488			
R: P	0.0000000	-0.0000027	0.0000018		
IP	-0.0000805	-0.0000037	0.0000011	0.0001404	
IR: P	0.0000000	0.0000011	-0.0000007	-0.0000399	0.0000266

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 33
D STUDY

D STUDY DESIGN NUMBER 001-011

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 7 3

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES					VARIANCE COMPONENTS FOR MEAN SCORES				
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	7	0.00555	0.00499	0.03887	1.0000	7	0.00555	0.00499
R: P	0.01764	1.0000	21	0.00084	0.00096	0.01764	1.0000	21	0.00084	0.00096
IP	0.01442	1.0000	7	0.00206	0.00846	0.01442	1.0000	7	0.00206	0.00846
IR: P	0.74815	1.0000	21	0.03563	0.00368	0.74815	1.0000	21	0.03563	0.00368

QFM = QUADRATIC FORM

		VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108		
EXPECTED OBSERVED SCORE	0.36025	0.60021	0.09070		
LOWER CASE DELTA	0.03769	0.19413	0.00667		
UPPER CASE DELTA	0.04408	0.20995	0.00807		
MEAN	0.01587	0.12599			
				GENERALIZABILITY COEFFICIENT = 0.89539 (8.55919)	
				PHI = 0.87978 (7.31787)	

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 34
D STUDY

D STUDY DESIGN NUMBER 001-011

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000015	0.0000249			
R: P	0.0000000	-0.0000014	0.0000009		
IP	-0.0000575	-0.0000019	0.0000005	0.0000716	
IR: P	0.0000000	0.0000005	-0.0000004	-0.0000204	0.0000136

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 35
D STUDY

D STUDY DESIGN NUMBER 001-012

OBJECT OF MEASUREMENT : I FACETS : P R: P
G STUDY POPULATION SIZE : INFINITE G STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY POPULATION SIZE : INFINITE D STUDY UNIVERSE SIZES : INFINITE INFINITE
D STUDY SAMPLE SIZE : 38 D STUDY SAMPLE SIZES : 9 3

VARIANCE COMPONENTS IN TERMS OF G STUDY UNIVERSE (OF ADMISSIBLE OBSERVATIONS) SIZES VARIANCE COMPONENTS IN TERMS OF D STUDY UNIVERSE (OF GENERALIZATION) SIZES

EFFECT	VARIANCE COMPONENTS FOR MEAN SCORES					VARIANCE COMPONENTS FOR MEAN SCORES				
	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS	VARIANCE COMPONENTS FOR SINGLE OBSERVATIONS	FINITE UNIVERSE CORRECTIONS	D STUDY SAMPLING FREQUENCIES	ESTIMATES	STANDARD ERRORS
I	0.32256	1.0000	1	0.32256	0.09108	0.32256	1.0000	1	0.32256	0.09108
P	0.03887	1.0000	9	0.00432	0.00388	0.03887	1.0000	9	0.00432	0.00388
R: P	0.01764	1.0000	27	0.00065	0.00074	0.01764	1.0000	27	0.00065	0.00074
IP	0.01442	1.0000	9	0.00160	0.00658	0.01442	1.0000	9	0.00160	0.00658
IR: P	0.74815	1.0000	27	0.02771	0.00287	0.74815	1.0000	27	0.02771	0.00287

QFM = QUADRATIC FORM

	VARIANCE	STANDARD DEVIATION	STANDARD ERROR OF VARIANCE	
UNIVERSE SCORE	0.32256	0.56795	0.09108	
EXPECTED OBSERVED SCORE	0.35188	0.59319	0.09074	
LOWER CASE DELTA	0.02931	0.17121	0.00519	GENERALIZABILITY COEFFICIENT = 0.91670 (11.00467)
UPPER CASE DELTA	0.03428	0.18516	0.00628	PHI = 0.90393 (9.40869)
MEAN	0.01423	0.11930		

NOTE: SIGNAL/NOISE RATIOS ARE IN PARENTHESES

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 36
D STUDY

D STUDY DESIGN NUMBER 001-012

VARIANCE - COVARIANCE MATRIX FOR ESTIMATED VARIANCE COMPONENTS (FOR MEAN SCORES) IN UNIVERSE OF GENERALIZATION (W)

	I	P	R: P	IP	IR: P
I	0.0082964				
P	0.0000012	0.0000151			
R: P	0.0000000	-0.0000008	0.0000006		
IP	-0.0000447	-0.0000011	0.0000003	0.0000433	
IR: P	0.0000000	0.0000003	-0.0000002	-0.0000123	0.0000082

GENOVA VERSION 3.1 FIRST SET OF D STUDY CONTROL CARDS PAGE 37
D STUDY

SUMMARY OF D STUDY RESULTS FOR SET OF CONTROL CARDS NO. 001

D. STUDY DESIGN NO	SAMPLE SIZES			VARIANCES								
	INDEX= UNIV. =	\$I INF.	P INF.	R INF.	UNIVERSE SCORE	EXPECTED OBSERVED SCORE	LOWER CASE DELTA	UPPER CASE DELTA	MEAN	GEN. COEF.	PHI	
001-001	38	3	1		0.32256	0.57675	0.25419	0.27303	0.03401	0.55927	0.54159	
001-002	38	5	1		0.32256	0.47508	0.15251	0.16382	0.02380	0.67897	0.66319	
001-003	38	7	1		0.32256	0.43150	0.10894	0.11701	0.01943	0.74754	0.73381	
001-004	38	9	1		0.32256	0.40729	0.08473	0.09101	0.01700	0.79197	0.77994	
001-005	38	3	2		0.32256	0.45206	0.12950	0.14539	0.02779	0.71354	0.68930	
001-006	38	5	2		0.32256	0.40026	0.07770	0.08724	0.02007	0.80588	0.78712	
001-007	38	7	2		0.32256	0.37806	0.05550	0.06231	0.01676	0.85320	0.83810	
001-008	38	9	2		0.32256	0.36573	0.04317	0.04846	0.01492	0.88197	0.86938	
001-009	38	3	3		0.32256	0.41050	0.08793	0.10285	0.02572	0.78579	0.75823	
001-010	38	5	3		0.32256	0.37532	0.05276	0.06171	0.01883	0.85943	0.83941	
001-011	38	7	3		0.32256	0.36025	0.03769	0.04408	0.01587	0.89539	0.87978	
001-012	38	9	3		0.32256	0.35188	0.02931	0.03428	0.01423	0.91670	0.90393	

GENOVA VERSION 3.1 CONTROL CARD INPUT LISTING PAGE 38

COLUMN 11111111122222222233333333334444444455555555666666667777777778
1234567890123456789012345678901234567890123456789012345678901234567890
FINISH



ประวัติผู้เขียนวิทยานิพนธ์

นางสิรินธร สิ้นจินดาวงศ์ เกิดวันที่ 21 มกราคม พ.ศ. 2515 ที่อำเภอสองพี่น้อง จังหวัดสุพรรณบุรี สำเร็จการศึกษาคณะศึกษาศาสตร์บัณฑิต สาขาวิชาวิทยาศาสตร์ทั่วไป จากวิทยาลัยครูพระนครศรีอยุธยา ในปีการศึกษา 2536 และการศึกษามหาบัณฑิต สาขาวิชาการวัดผลการศึกษ จากมหาวิทยาลัยศรีนครินทรวิโรฒ ในปีการศึกษา 2543 และเข้าศึกษาต่อหลักสูตรครุศาสตรดุษฎีบัณฑิต สาขาวิชาการวัดและประเมินผลการศึกษา จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2549 ปัจจุบันเป็นอาจารย์ประจำสำนักงานแผนและประกันคุณภาพ มหาวิทยาลัยศรีปทุม

ทุนนำเสนอผลงานวิทยานิพนธ์ในต่างประเทศ

ได้รับทุนนำเสนอผลงานวิทยานิพนธ์ในต่างประเทศ จากมหาวิทยาลัยศรีปทุม ในหัวข้อ การประชุม Evaluation 2009: Context and Evaluation จัดโดย The American Evaluation Association (AEA) จัดขึ้นที่ The world class Rosen Shingle Creek Resort, Orlando, Florida ประเทศสหรัฐอเมริกา ระหว่างวันที่ 11 - 14 พฤศจิกายน พ.ศ. 2552

ผลงานวิจัยระหว่างศึกษา

นางลักษณ์ วิรัชชัย, สิรินธร สิ้นจินดาวงศ์, ภัทราวดี มากมี และ นัทธี เชียงชนะ. (2551).

การสังเคราะห์งานวิจัยที่นำเสนอในการประชุมวิชาการ “เปิดขอบฟ้าคุณธรรมจริยธรรม”.

กรุงเทพฯ: ศูนย์ส่งเสริมและพัฒนาพลังแผ่นดินเชิงคุณธรรม (ศูนย์คุณธรรม) สำนักงานบริหารและพัฒนาองค์ความรู้ (องค์การมหาชน).

นางลักษณ์ วิรัชชัยและคณะ. (2551). รายงานการสังเคราะห์งานวิจัยเกี่ยวกับคุณภาพการศึกษาไทย:

การวิเคราะห์ห่อภิมาณ (Meta-analysis). กรุงเทพฯ: สำนักงานเลขาธิการสภาการศึกษา (สกศ.). (ผู้ร่วมวิจัย)

ขวัญเมือง แก้วดำเกิง และคณะ. (2552). รายงานการประเมินแผนงานสุขภาวะครอบครัว.

กรุงเทพฯ: สำนักงานกองทุนสนับสนุนการสร้างเสริมสุขภาพ (สสส.). (ผู้ร่วมวิจัย)

สิรินธร สิ้นจินดาวงศ์ และคณะ. (2552). การวิเคราะห์คุณภาพของแบบทดสอบการศึกษา

ระดับชาติขั้นพื้นฐาน (O-NET) แบบทดสอบความถนัดทั่วไป (GAT) และความถนัดทางวิชาการและวิชาชีพ (PAT). สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).