



รายงานวิจัยฉบับสมบูรณ์

การสังเคราะห์งานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียน
ในศตวรรษที่ 21 ระดับมัธยมศึกษา

A SYNTHESIS OF RESEARCH ON THE 21ST CENTURY COMPETENCIES
MEASUREMENT TOOLS DEVELOPMENT
FOR SECONDARY SCHOOL STUDENTS

ผู้ช่วยศาสตราจารย์ ดร.สรัญญา จันทร์ชูสกุล
ผู้ช่วยศาสตราจารย์ ดร.ทัศนศิริรินทร์ สว่างบุญ

งานวิจัยนี้ได้รับเงินอุดหนุนการวิจัยจาก
สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)
ประจำปีงบประมาณ พ.ศ. 2562

ชื่อโครงการ การสังเคราะห์งานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

ประจำปีงบประมาณ พ.ศ. 2562

จำนวนเงินที่ได้รับการสนับสนุน 150,000 บาท

ระยะเวลาทำการวิจัย 7 เดือน

ตั้งแต่ 1 สิงหาคม 2562 ถึง 28 กุมภาพันธ์ 2563

ผู้ช่วยศาสตราจารย์ ดร.สรัญญา จันทร์ชูสกุล มหาวิทยาลัยศิลปากร และ ผู้ช่วยศาสตราจารย์ ดร.ทัศนศิริรินทร์ สว่างบุญ มหาวิทยาลัยมหาสารคาม

บทคัดย่อ

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อ 1) ประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา 2) ศึกษาคุณลักษณะของงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา และ 3) สังเคราะห์องค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา งานวิจัยที่นำมาสังเคราะห์มีจำนวน 92 เล่ม เครื่องมือที่ใช้ประกอบด้วย แบบประเมินคุณภาพงานวิจัย และแบบบันทึกคุณลักษณะงานวิจัย เครื่องมือทั้งสองฉบับมีค่าความสอดคล้องระหว่างผู้ประเมินอยู่ระหว่าง 0.88-0.94 และ 0.96-0.99 ตามลำดับ วิเคราะห์ข้อมูลด้วยสถิติเชิงบรรยาย การวิเคราะห์ความแปรปรวนทางเดียว และการวิเคราะห์เนื้อหา ผลการวิจัยพบว่า

1. ผลการประเมินคุณภาพงานวิจัย พบว่า งานวิจัยส่วนใหญ่มีคุณภาพอยู่ในระดับดี (ร้อยละ 82.61)

2. ผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา พบว่า

2.1 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความเที่ยง มีจำนวน 11 ตัวแปร ได้แก่ ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, เพศของผู้วิจัย, ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, การตรวจให้คะแนน, จำนวนข้อคำถามในแต่ละฉบับ และวิธีการเลือกตัวอย่าง

2.2 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความตรง มีเพียงตัวแปรเดียวคือ เนื้อหาที่เกี่ยวข้อง

2.3 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความยาก มีจำนวน 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทงานวิจัย, การสร้างเครื่องมือวัด, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์ และวิธีการเลือกตัวอย่าง

2.4 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานอำนาจจำแนก มีจำนวน 10 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, เพศของผู้วิจัย, ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนข้อคำถามในแต่ละฉบับ และจำนวนผู้สอบในแต่ละฉบับ

3. ผลการสังเคราะห์องค์ความรู้ พบว่า ขั้นตอนการพัฒนาเครื่องมือในแต่ละสมรรถนะมีความคล้ายคลึงกัน แต่มีความแตกต่างของประเภทเครื่องมือที่ใช้วัด และเกณฑ์การให้คะแนน โดยแบ่งประเภทเครื่องมือได้ 4 กลุ่มคือ 1) มาตรฐานค่า 2) แบบวัดเชิงสถานการณ์ 3) แบบสอบหลายตัวเลือก และ 4) แบบสอบอัตนัย

คำสำคัญ : เครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21, การวิเคราะห์ห่อภิมาณ, มัธยมศึกษา

Research Title: A synthesis of research on the 21st century competencies measurement tools development for secondary school students

Researcher: Asst. Prof. Saranya Chanchusakun, Ph.D., Silpakorn University
Asst. Prof. Tatsirin Sawangboon, Ph.D., Mahasarakram University

Abstract

This research aimed to 1) evaluate the researches quality on the 21st century competencies measurement tools development for secondary school students, 2) study characteristics of research affecting the standard indices of the 21st century competencies measurement tools quality of the secondary school students and 3) synthesize knowledge about the development of tools for measuring the 21st century competencies of secondary school students. The synthesized researches were 92 research papers. The research instruments were research quality evaluative forms and research characteristics coding forms. Both of the instruments had inter-rater agreement ranged from 0.88-0.94 and 0.96-0.99 respectively. The data analysis included descriptive statistics, one-way ANOVA and content analysis. The findings were as follows:

1. The result of evaluate the researches quality was found that most theses were of good quality (82.61%)

2. The result of research characteristic affecting the standard indices of the 21st century competencies measurement tools quality of the secondary school students found that

2.1 There were 10 research characteristic variables affecting the reliability standard indices which were types of institutions that produce research, institutions that produce research, fields of research, researcher genders, regions, student levels, types of measurement tool, number of choices or criteria, scoring, number of items in subtest and sampling designs.

2.2 There was only one research characteristic variables affecting the validity standard indices which was related content.

2.3 There were 6 research characteristic variables affecting the difficulty standard indices which were year of research published, types of research, measurement tool developing, types of measurement tool and number of choices or criteria and sampling designs.

2.4 There were 10 research characteristic variables affecting the discrimination standard indices which were year of research published, types of institutions that produce research, institutions that produce research, fields of research, researcher genders, regions, student levels, types of measurement tools, number of items in subtest and sample sizes.

3. The synthesis knowledge found that the steps of measurement tools development in each competency were similar but there were differences in the types and scoring criterion of measuring tools. There were 4 groups of tools classified as follow 1) rating scale 2) situation test 3) multiple choice test and 4) subjective test.

Keywords : 21st century competencies measurement tools, meta-analysis, secondary education

กิตติกรรมประกาศ

โครงการวิจัยการสังเคราะห์งานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา สำเร็จลงด้วยดีด้วยการสนับสนุนจากคณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร และคณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม ในการทำวิจัยครั้งนี้

คณะผู้วิจัยขอขอบพระคุณผู้เชี่ยวชาญ ได้แก่ รองศาสตราจารย์ ดร.กมลวรรณ ตังชันกานนท์, รองศาสตราจารย์ ดร.ณัฐภรณ์ หลาวทอง, รองศาสตราจารย์ ดร.ไชยยศ ไพวิทยศิริธรรม, ผู้ช่วยศาสตราจารย์ ดร.พินดา วราสุนันท์ และผู้ช่วยศาสตราจารย์ ดร.วราพร เอราวรรณ ที่ได้เสียสละเวลาและให้ข้อเสนอแนะเพื่อการปรับปรุงเครื่องมือให้ได้เครื่องมือที่มีคุณภาพ

ขอขอบคุณสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ที่ช่วยสนับสนุนงบประมาณการวิจัยในครั้งนี้

ขอขอบพระคุณคณาจารย์ทุกท่านที่ได้อบรมสั่งสอน ประสิทธิ์ประสาทวิชาให้แก่คณะผู้วิจัย และขอขอบคุณกำลังใจจากครอบครัวที่ทำให้งานในครั้งนี้สำเร็จลงด้วยดี

คณะผู้วิจัย
กุมภาพันธ์ 2563

สารบัญ

	หน้า
บทคัดย่อ.....	ก
Abstract.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ฉ
สารบัญภาพ.....	ช
บทที่ 1 บทนำ.....	1
ความเป็นมาและความสำคัญของปัญหา.....	1
คำถามวิจัย.....	3
วัตถุประสงค์การวิจัย.....	3
ขอบเขตการวิจัย.....	3
นิยามศัพท์เฉพาะ.....	4
ประโยชน์ที่ได้รับจากการวิจัย.....	5
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง.....	6
ตอนที่ 1 สมรรถนะของผู้เรียนในศตวรรษที่ 21.....	6
ตอนที่ 2 การพัฒนาเครื่องมือวัด.....	22
ตอนที่ 3 การสังเคราะห์งานวิจัย.....	44
ตอนที่ 4 กรอบแนวคิดในการวิจัย.....	58
บทที่ 3 วิธีดำเนินการวิจัย.....	60
ประชากรและตัวอย่าง.....	60
ตัวแปรที่ใช้ในการวิจัย.....	61
เครื่องมือที่ใช้ในการวิจัย.....	68
การเก็บรวบรวมข้อมูล.....	78
การวิเคราะห์ข้อมูลและสถิติที่ใช้.....	78
บทที่ 4 ผลการวิเคราะห์ข้อมูล.....	79
ตอนที่ 1 ผลการประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของ ผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา.....	79
ตอนที่ 2 ผลการวิเคราะห์ข้อมูลเบื้องต้นงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะ ของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษาที่นำมาสังเคราะห์.....	82
ตอนที่ 3 ผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพ เครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา....	102
ตอนที่ 4 ผลสังเคราะห์องค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของ ผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา.....	136
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ.....	143

	หน้า
สรุปผลการวิจัย.....	142
อภิปรายผลการวิจัย.....	146
ข้อเสนอแนะในการวิจัย.....	155
รายการอ้างอิง.....	158
ภาคผนวก ก รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์.....	162
ภาคผนวก ข รายชื่อผู้เชี่ยวชาญ.....	173
ภาคผนวก ค เครื่องมือที่ใช้ในการวิจัย.....	175
ภาคผนวก ง ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือ วัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา.....	201

สารบัญตาราง

ตารางที่	หน้า
2.1 ผลการสังเคราะห์สมรรถนะของผู้เรียนในศตวรรษที่ 21.....	17
2.2 การเปรียบเทียบวิธีการสังเคราะห์งานวิจัย.....	46
3.1 รายละเอียดตัวแปรคุณลักษณะงานวิจัยจำแนกตามการนำไปใช้วิเคราะห์ข้อมูล.....	67
3.2 โครงสร้างของแบบประเมินคุณภาพงานวิจัย.....	69
3.3 ผลการปรับปรุงแบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพตาม ข้อเสนอแนะของผู้เชี่ยวชาญ.....	69
3.4 ผลการตรวจสอบความสอดคล้องระหว่างผู้ประเมินและภายในตัวผู้ประเมินจากการ ประเมินคุณภาพงานวิจัย.....	75
3.5 ผลการปรับปรุงแบบบันทึกคุณลักษณะงานวิจัย และคู่มือลงรหัสตามข้อเสนอแนะของ ผู้เชี่ยวชาญ.....	76
3.6 ผลการตรวจสอบความสอดคล้องระหว่างผู้ประเมินและภายในตัวผู้ประเมินจากการ บันทึกคุณลักษณะงานวิจัย.....	77
4.1 ระดับคุณภาพงานวิจัยจำแนกตามประเภทสถาบันที่ผลิตงานวิจัยและประเภทงานวิจัย.	80
4.2 ผลการประเมินคุณภาพงานวิจัย.....	81
4.3 ข้อมูลพื้นฐานคุณลักษณะงานวิจัยระดับเล่งงานวิจัยที่เป็นตัวแปรจัดประเภท.....	87
4.4 ข้อมูลพื้นฐานของตัวแปรคุณลักษณะงานวิจัยที่เป็นตัวแปรต่อเนื่อง.....	98
4.5 ข้อมูลเบื้องต้นของการทดลองใช้เครื่องมือ.....	100
4.6 ผลการวิเคราะห์ข้อมูลเบื้องต้นของค่าดัชนีมาตรฐาน.....	103
4.7 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงจำแนกตามตัวแปร คุณลักษณะงานวิจัย.....	108
4.8 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรงจำแนกตามตัวแปร คุณลักษณะงานวิจัย.....	114
4.9 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยากจำแนกตามตัวแปร คุณลักษณะงานวิจัย.....	120
4.10 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกจำแนกตามตัว แปรคุณลักษณะงานวิจัย.....	128
4.11 สรุปคุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะ ของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา.....	134

สารบัญภาพ

ภาพที่	หน้า
2.1 กรอบแนวคิดหลักของสมรรถนะของผู้เรียนในศตวรรษที่ 21 ของสิงคโปร์.....	13
2.2 กรอบแนวคิดในการวิจัย.....	59

บทที่ 1 บทนำ

ความเป็นมาและความสำคัญของปัญหา

การเข้าสู่ศตวรรษที่ 21 และการเปลี่ยนแปลงแบบก้าวกระโดดทางเทคโนโลยี ทำให้เกิดความตื่นตัวของนักการศึกษา นักวิชาการ รวมถึงองค์กรเอกชนในการเสนอแนวคิดที่เกี่ยวกับกรอบการเรียนรู้ในศตวรรษที่ 21 เพื่อวางเป้าหมายและเตรียมความพร้อมให้ผู้เรียนมีสมรรถนะที่จำเป็นในศตวรรษที่ 21 โดยเริ่มในปี ค.ศ.2003 กลุ่ม METIRI (METIRI, 2003) ประเทศสหรัฐอเมริกา ได้ใช้เวลาในการพัฒนาและเสนอกรอบทักษะที่จำเป็นในศตวรรษที่ 21 ที่ชื่อว่า “enGauge 21st century skills” ประกอบด้วยทักษะที่จำเป็น 4 ด้าน สำหรับในยุคเทคโนโลยีสารสนเทศ ได้แก่ ความรู้ความสามารถในยุคดิจิทัล (digital-age literacy) การคิดริเริ่ม (inventive thinking) การสื่อสารอย่างมีประสิทธิภาพ (effective communication) และการเพิ่มผลผลิตในระดับสูง (high productivity) ในเวลาต่อมา Partnership for 21st Century Learning หรือ P21 (2019) ได้พัฒนารอบแนวคิดการเรียนรู้ในศตวรรษที่ 21 โดยครูผู้เชี่ยวชาญทางการศึกษา และผู้นำทางภาคธุรกิจในประเทศสหรัฐอเมริกา เพื่อที่กำหนดและแสดงให้เห็นถึงความรู้และทักษะที่ผู้เรียนจำเป็นต้องมีเพื่อความสำเร็จในอาชีพและการทำงาน ด้วยเหตุนี้ทำให้ทั่วโลกเกิดการตื่นตัว และได้มีการกำหนดกรอบสมรรถนะเพื่อเป็นเป้าหมายในการพัฒนาผู้เรียนที่เหมาะสมกับบริบทในประเทศของตน

สำหรับประเทศไทยได้มีการบรรจุเรื่องการพัฒนาผู้เรียนให้มีคุณลักษณะและทักษะการเรียนรู้ ในศตวรรษที่ 21 ไว้ใน แผนการศึกษาแห่งชาติ พ.ศ. 2560-2579 (สำนักงานเลขาธิการสภาการศึกษา, 2560) ดังที่ได้กำหนดวิสัยทัศน์ไว้ว่า “คนไทยทุกคนได้รับการศึกษาและเรียนรู้ตลอดชีวิตอย่างมีคุณภาพ ดำรงชีวิตอย่างเป็นสุข สอดคล้องกับหลักปรัชญาของเศรษฐกิจพอเพียง และการเปลี่ยนแปลงของโลกศตวรรษที่ 21” โดยมีวัตถุประสงค์ในการจัดการศึกษา 4 ประการ คือ 1) เพื่อพัฒนาระบบและกระบวนการ จัดการศึกษาที่มีคุณภาพและมีประสิทธิภาพ 2) เพื่อพัฒนาคนไทยให้เป็นพลเมืองดี มีคุณลักษณะ ทักษะและสมรรถนะที่ สอดคล้องกับบทบาทบัญญัติของรัฐธรรมนูญแห่งราชอาณาจักรไทย พระราชบัญญัติการศึกษาแห่งชาติและ ยุทธศาสตร์ชาติ 3) เพื่อพัฒนาสังคมไทยให้เป็นสังคม แห่งการเรียนรู้ และคุณธรรม จริยธรรม รู้รักสามัคคี และร่วมมือผนึกกำลังมุ่งสู่การพัฒนาประเทศ อย่างยั่งยืน ตามหลักปรัชญาของเศรษฐกิจพอเพียง และ 4) เพื่อนำประเทศไทยก้าวข้ามกับดัก ประเทศที่มีรายได้ปานกลาง และความเหลื่อมล้ำภายในประเทศลดลง อย่างไรก็ตาม เพื่อให้การดำเนินบรรลุวิสัยทัศน์และจุดมุ่งหมายในการจัดการศึกษาดังกล่าวข้างต้น แผนการ ศึกษาแห่งชาติได้วางเป้าหมายไว้ 2 ด้าน คือ เป้าหมายด้านผู้เรียน (learner aspirations) และเป้าหมาย ของการจัดการศึกษา (aspirations) รวมถึงได้กำหนดคุณลักษณะและทักษะการเรียนรู้ ในศตวรรษที่ 21 ประกอบด้วยทักษะที่เรียกตามคำย่อว่า 3Rs + 8Cs ประกอบด้วย ทักษะและคุณลักษณะต่อไปนี้ 3Rs ได้แก่ อ่านออก (Reading) เขียนได้ (WRiting) และ คิดเลขเป็น (ARithmetics) 8Cs ประกอบด้วย 1) ทักษะด้าน การคิดอย่างมีวิจารณญาณ และทักษะในการแก้ปัญหา (critical thinking and problem solving) 2) ทักษะด้านการสร้างสรรค์และนวัตกรรม (creativity and innovation) 3) ทักษะด้านความเข้าใจต่าง วัฒนธรรม ต่างกระบวนทัศน์ (cross-cultural understanding) 4) ทักษะด้านการทำงานร่วมกัน การทำงาน เป็นทีม และภาวะผู้นำ (collaboration, teamwork and leadership) 5) ทักษะด้านการสื่อสาร สารสนเทศ และการรู้เท่าทันสื่อ (communications, information and media literacy) 6) ทักษะด้านคอมพิวเตอร์ และ เทคโนโลยีสารสนเทศและการสื่อสาร (computing and ICT literacy) 7) ทักษะอาชีพ และทักษะ

การเรียนรู้ (career and learning skills) และ 8) ความมีเมตตา กรุณา มีวินัย คุณธรรม จริยธรรม (compassion)

ด้วยการเล็งเห็นความสำคัญของการพัฒนาผู้เรียนให้มีคุณลักษณะและทักษะการเรียนรู้ ในศตวรรษที่ 21 จึงก่อให้เกิดงานวิจัยที่เกี่ยวข้องในหลากหลายมิติ หนึ่งในนั้นคือ การพัฒนาเครื่องมือในการวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ซึ่งหากมีการรวบรวมและสังเคราะห์งานวิจัยที่พัฒนาเครื่องมือวัดดังกล่าวจะทำให้ได้สารสนเทศที่เป็นประโยชน์ต่อครู บุคลากรทางการศึกษา และนักวิชาการในการเลือกเครื่องมือวัดที่ผ่านการรวบรวมอย่างเป็นระบบจากการวิจัยไปใช้ได้อย่างเหมาะสม ในขณะที่นักวิจัยที่สนใจพัฒนาเครื่องมือวัดจะได้รับสารสนเทศเกี่ยวกับคุณลักษณะที่สำคัญที่ส่งผลกระทบต่อการพัฒนาเครื่องมือวัดสมรรถนะที่มีคุณภาพ ซึ่งจะช่วยให้มีข้อมูลข้อที่ควรพิจารณาและให้ความสนใจเพื่อการพัฒนาเครื่องมือวัดที่มีคุณภาพ รวมถึงยังได้แนวคิดในการพัฒนาเครื่องมือวัดในอนาคต

ในการสังเคราะห์งานวิจัย การวิเคราะห์อภิमानเป็นการสังเคราะห์งานวิจัยเชิงปริมาณแบบหนึ่ง ที่นักวิจัยนำงานวิจัยซึ่งศึกษาปัญหาเดียวกันหลาย ๆ เรื่องมาศึกษาวิเคราะห์ด้วยวิธีการทางสถิติ เพื่อให้ได้ข้อสรุปผลการวิจัยที่มีความลุ่มลึกกว่าผลการวิจัยที่นำมาสังเคราะห์แต่ละเรื่อง ข้อมูลสำหรับการวิเคราะห์อภิमानประกอบด้วยข้อมูล 2 ส่วน ส่วนแรก คือ ดัชนีมาตรฐาน (standard index) ข้อมูลในส่วนที่สองเป็นข้อมูลเกี่ยวกับคุณลักษณะงานวิจัยที่นำมาสังเคราะห์ ได้แก่ ด้านการพิมพ์และผู้วิจัย ด้านเนื้อหาสาระงานวิจัย ด้านวิธีวิทยาการวิจัย และด้านคุณภาพงานวิจัย โดยการวิเคราะห์ข้อมูลในการวิเคราะห์อภิमानมีจุดมุ่งหมายที่สำคัญ 2 ประการคือ การสังเคราะห์ให้ได้ข้อสรุปเกี่ยวกับดัชนีมาตรฐาน และการวิเคราะห์เพื่ออธิบายความแตกต่างของดัชนีมาตรฐานด้วยตัวแปรคุณลักษณะงานวิจัย หรือการวิเคราะห์เพื่อตรวจสอบความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรคุณลักษณะงานวิจัยในฐานะที่เป็นตัวแปรปรับหรือตัวแปรกำกับ (moderator) ที่มีอิทธิพลทำให้ดัชนีมาตรฐานที่เป็นผลการวิจัยแตกต่างกัน (นงลักษณ์ วิรัชชัย, 2542; สำนักเลขาธิการสภาการศึกษา, 2552; Koricheva, Gurevitch & Mengersen, 2013) อย่างไรก็ตาม ได้มีผู้นำวิธีการวิเคราะห์อภิमानมาประยุกต์ใช้ในการสังเคราะห์งานวิจัยเกี่ยวกับการพัฒนาแบบวัด เช่น งานวิจัยของ รัชนิกุล ภิโยภานูวัฒน์ (2536) ได้วิเคราะห์อภิमानงานวิจัยประเภทวิทยานิพนธ์เกี่ยวกับการสร้างและพัฒนาแบบสอบและแบบวัด ด้วยวิธีของ Glass ในขณะที่ ไชยยศ ไพวิทยศิริธรรม (2544) ได้ทำการสังเคราะห์งานวิจัยเกี่ยวกับคุณภาพของแบบสอบหลายตัวเลือก ประกอบด้วย ค่าความเที่ยง ค่าความตรง ค่าความยาก และค่าอำนาจจำแนก ที่มีรูปแบบของแบบสอบ บริบทของแบบสอบ วิธีการตรวจให้คะแนน และประเภทของแบบสอบต่างกัน ด้วยวิธีการวิเคราะห์อภิमानตามแนวคิดของ Hunter รวมถึงงานวิจัยของ อภาพรรณ ประทุมไทย (2553) ได้ดำเนินการวิเคราะห์อภิमानคุณภาพของแบบวัดความถนัด จากงานวิจัยทั้ง 3 เรื่อง ทำให้ได้ข้อสรุปเกี่ยวกับปัจจัยที่ส่งผลกระทบต่อคุณภาพเครื่องมือวัดทั้งในมิติความเที่ยง ความตรง ความยาก และอำนาจจำแนก สารสนเทศดังกล่าวจะเป็นประโยชน์ต่อการพัฒนาเครื่องมือวัดต่อไปในอนาคต นอกจากนี้หากมีการดำเนินการสังเคราะห์งานวิจัยวิธีการเชิงคุณภาพด้วยการวิเคราะห์เนื้อหา จะทำให้ได้สารสนเทศและข้อค้นพบที่มีความลุ่มลึก ครอบคลุมในมิติทั้งเชิงปริมาณและเชิงคุณภาพทำให้ต้องรู้เกี่ยวกับการพัฒนาเครื่องมือวัดที่เป็นประโยชน์มากยิ่งขึ้น

จากความสำคัญดังกล่าวข้างต้น ผู้วิจัยจึงสนใจสังเคราะห์งานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 สำหรับผู้เรียนในระดับมัธยมศึกษา โดยทำการสังเคราะห์ในเชิงปริมาณด้วยวิธีการวิเคราะห์อภิमान และการสังเคราะห์เชิงคุณภาพด้วยวิธีการวิเคราะห์เนื้อหา

คำถามการวิจัย

1. คุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา อยู่ในระดับใด
2. คุณลักษณะใดของงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา
3. องค์ความรู้ที่ได้จากการสังเคราะห์งานวิจัยเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษาเป็นอย่างไร

วัตถุประสงค์การวิจัย

1. เพื่อประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา
2. เพื่อศึกษาคุณลักษณะของงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา
3. เพื่อสังเคราะห์องค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

ขอบเขตการวิจัย

ในการสังเคราะห์งานวิจัยเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ในครั้งนี้ผู้วิจัยเลือกวิธีการสังเคราะห์ด้วยการวิเคราะห์เนื้อหาและการสังเคราะห์เชิงเนื้อหา เพื่อให้ครอบคลุมทั้งในมิติปริมาณและคุณภาพ โดยงานวิจัยที่คัดเลือกมาใช้ในการสังเคราะห์เป็นงานวิจัยประเภทวิทยานิพนธ์ระดับบัณฑิตศึกษา ทั้งระดับมหาบัณฑิตและดุษฎีบัณฑิต และงานวิจัยของสถาบันต่าง ๆ โดยไม่นับรวมสารนิพนธ์หรือการค้นคว้าอิสระ และเป็นงานวิจัยที่มีการตีพิมพ์เผยแพร่ระหว่างปี พ.ศ. 2546-2562 เนื่องจากเป็นปีที่เริ่มมีการกำหนดกรอบแนวคิดสำหรับทักษะแห่งศตวรรษที่ 21 ภายใต้ความร่วมมือระหว่าง ห้องวิจัยด้านการศึกษาระดับภูมิภาคกลางตอนเหนือ ในประเทศสหรัฐอเมริกา (North Central Regional Educational Laboratory : NCREL) และกลุ่ม Metiri ในชื่อของกรอบแนวคิด “enGauge” (Metiri Group, 2003) ผู้วิจัยดำเนินการสืบค้นและคัดเลือกงานวิจัยที่สืบค้นได้จากห้องสมุดสถาบันอุดมศึกษา และฐานข้อมูลออนไลน์ เช่น เว็บไซต์โครงการเครือข่ายห้องสมุดในประเทศไทย หรือ ThaiLIS (www.thailis.or.th/tdc/), ศูนย์ข้อมูลการวิจัย Digital "วช." (<https://dric.nrct.go.th/Index>) และเว็บไซต์ห้องสมุดของสถาบันอุดมศึกษาต่าง ๆ

ทั้งนี้ สมรรถนะในศตวรรษที่ 21 ในการวิจัยครั้งนี้ได้มาจากการสังเคราะห์จากเอกสารที่นำเสนอเกี่ยวกับสมรรถนะในศตวรรษที่ 21 จากประเทศต่าง ๆ ผลการสังเคราะห์พบสมรรถนะที่สำคัญแบ่งออกเป็น 4 ด้าน คือ 1) ทักษะและความรู้ในเนื้อหาขั้นสูง ประกอบด้วย จิตสำนึกต่อโลก (global awareness) ความรู้เรื่องพลเมือง (civic literacy) 2) ทักษะการคิดขั้นสูง ประกอบด้วย การคิดอย่างมีวิจารณญาณ (critical thinking) การแก้ปัญหา (solving problems) และการสร้างสรรค์และนวัตกรรม (creativity and innovation) 3) ทักษะต่อตนเองและสังคม ประกอบด้วย การสื่อสาร (communications literacy/skill) ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (collaboration, teamwork) ภาวะผู้นำ (leadership) และความรับผิดชอบ (responsibility) และ 4) ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี

ประกอบด้วย การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy) และการรู้เท่าทันเทคโนโลยี (technology literacy)

นิยามศัพท์เฉพาะ

เครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 หมายถึง เครื่องมือประเภทแบบสอบถาม แบบวัดต่าง ๆ ที่ใช้สำหรับวัดสมรรถนะเกี่ยวกับจิตสำนึกต่อโลก ความรู้เรื่องพลเมือง การคิดอย่างมีวิจารณญาณ การแก้ปัญหา สร้างสรรค์และนวัตกรรม การสื่อสาร ทักษะด้านการทำงานเป็นทีม ความร่วมมือ ภาวะผู้นำ ความรับผิดชอบ การรู้เท่าทันสื่อและสารสนเทศ และการรู้เท่าทันเทคโนโลยี ของผู้เรียนในระดับมัธยมศึกษา ที่ผ่านการพัฒนาเครื่องมือวัดด้วยกระบวนการวิจัยอย่างเป็นระบบ

การวิเคราะห์ห่อภิมาณ หมายถึง วิธีการวิจัยที่ใช้ในการสังเคราะห์งานวิจัยเชิงปริมาณเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา โดยใช้เทคนิควิธีการทางสถิติในการวิเคราะห์ข้อมูลที่เป็นผลการวิจัย ซึ่งวัดในรูปของค่าดัชนีมาตรฐานแต่ละค่าเป็นข้อมูลในการวิเคราะห์และใช้ข้อมูลคุณลักษณะ ประกอบด้วย ด้านการพิมพ์และผู้วิจัย ด้านเนื้อหาสาระงานวิจัย ด้านวิธีวิทยาการวิจัย และด้านคุณภาพงานวิจัยมาเป็นตัวแปรปรับในการวิเคราะห์ห่อภิมาณทำให้ดัชนีมาตรฐานที่เป็นผลการวิจัยแตกต่างกัน เพื่อให้ได้ข้อสรุปที่กว้างขวางลุ่มลึก

ค่าดัชนีมาตรฐาน หมายถึง ค่าสถิติซึ่งแสดงค่าคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ที่ได้จากงานวิจัยแต่ละเล่ม ในการวิจัยครั้งนี้ใช้ค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัด 4 ประเภท ประกอบด้วย ค่าดัชนีความเที่ยง ค่าดัชนีความตรง ค่าดัชนีความยาก และค่าดัชนีอำนาจจำแนก โดยการวิจัยครั้งนี้มีการกำหนดนิยามของค่าดัชนีมาตรฐานแต่ละประเภทไว้ดังนี้

ค่าดัชนีความเที่ยง หมายถึง ดัชนีที่บ่งบอกถึงความคงเส้นคงวาของผลที่ได้จากการวัด โดยพิจารณาการวิเคราะห์ค่าสัมประสิทธิ์ความเที่ยง ซึ่งงานวิจัยแต่ละเรื่องอาจใช้วิธีการประมาณค่าสัมประสิทธิ์ความเที่ยงที่แตกต่างกัน

ค่าดัชนีความตรง หมายถึง ดัชนีที่บ่งบอกถึงเครื่องมือวัดแต่ละฉบับสามารถวัดได้ในสิ่งที่ต้องการจะวัดได้ โดยในการวิจัยครั้งนี้พิจารณาเฉพาะดัชนีความตรงตามเกณฑ์สัมพัทธ์ 2 ประเภท ได้แก่ ความตรงตามสภาพ (concurrent validity) และความตรงเชิงทำนาย (predictive validity) เท่านั้น

ค่าดัชนีความยาก หมายถึง ค่าดัชนีที่บ่งบอกถึงสัดส่วนของจำนวนผู้สอบที่ตอบข้อสอบข้อนั้นได้ถูกต้อง ในการวิจัยครั้งนี้พิจารณาค่าดัชนีความยากในรายฉบับ โดยการนำค่าดัชนีความยากรายข้อมาหาค่าเฉลี่ย และศึกษาเฉพาะการวิเคราะห์ข้อสอบด้วยทฤษฎีการทดสอบแบบดั้งเดิมเท่านั้น

ค่าดัชนีอำนาจจำแนก หมายถึง ค่าดัชนีที่บ่งบอกถึงระดับความสามารถของข้อสอบในการจำแนกผู้สอบที่มีความสามารถแตกต่างกันออกจากกันได้ ในการวิจัยครั้งนี้พิจารณาค่าดัชนีอำนาจจำแนกในรายฉบับ โดยการนำค่าดัชนีอำนาจจำแนกรายข้อมาหาค่าเฉลี่ย และศึกษาเฉพาะการวิเคราะห์ข้อสอบด้วยทฤษฎีการทดสอบแบบดั้งเดิมเท่านั้น

คุณลักษณะงานวิจัย หมายถึง ตัวแปรซึ่งเป็นคุณลักษณะของงานวิจัย แบ่งออกเป็น 4 ด้าน คือ ด้านการพิมพ์และผู้วิจัย ด้านเนื้อหาสาระงานวิจัย ด้านวิธีวิทยาการวิจัย และด้านคุณภาพงานวิจัย ซึ่งใช้เพื่ออธิบายอิทธิพลที่ทำให้ดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 แตกต่างกัน

องค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 หมายถึง ข้อค้นพบสำคัญที่ได้จากการสังเคราะห์งานวิจัยที่เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนใน

ศตวรรษที่ 21 ระดับมัธยมศึกษา ด้วยการวิเคราะห์เนื้อหาในมิติขององค์ประกอบของเครื่องมือวัด ลักษณะของเครื่องมือวัด และแนวทางการพัฒนาเครื่องมือวัด

ประโยชน์ที่ได้รับจากการวิจัย

1. ได้ข้อมูลสารสนเทศเกี่ยวกับคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ซึ่งจะเป็นประโยชน์ต่อนิสิตนักศึกษาระดับบัณฑิตศึกษาที่อยู่ระหว่างวางแผนหรือจัดทำวิทยานิพนธ์ รวมถึงอาจารย์ที่ปรึกษาวิทยานิพนธ์ และนักวิจัย ซึ่งสารสนเทศเกี่ยวกับคุณภาพงานวิจัยจะทำให้ นิสิตนักศึกษา อาจารย์ที่ปรึกษา และนักวิจัยได้ทราบว่าในการทำงานวิจัยที่ผ่านมา ประเด็นใดที่ยังเป็นจุดอ่อนหรือเป็นประเด็นที่ควรให้ความสำคัญในการควบคุมงานวิจัยให้มีคุณภาพ

2. ได้ข้อสรุปเกี่ยวกับปัจจัยที่ส่งผลต่อคุณภาพของเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ทั้งค่าความตรง ค่าความเที่ยง ค่าความยาก และค่าอำนาจจำแนก ซึ่งเป็นข้อค้นพบที่เป็นประโยชน์ต่อครู นักวิชาการ และนักวิจัยในการสร้างและพัฒนาเครื่องมือวัดสมรรถนะให้มีคุณภาพ และสามารถนำไปใช้ทดสอบได้อย่างน่าเชื่อถือ

3. ได้องค์ความรู้ของการพัฒนาเครื่องมือวัดสมรรถนะที่เป็นปัจจุบัน ทั้งในมิติของเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ที่มีการพัฒนาไว้แล้ว ทำให้สามารถเลือกหรือประยุกต์ใช้ได้อย่างเหมาะสม รวมถึงมุมมองของการพัฒนาเครื่องมือที่น่าสนใจ และประเด็นการพัฒนาเครื่องมือวัดสมรรถนะที่ควรมีดำเนินการต่อไปในอนาคต

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการดำเนินการวิจัยเรื่อง การสังเคราะห์งานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ผู้วิจัยได้ดำเนินการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง เพื่อเป็นประโยชน์ในการดำเนินการวิจัย โดยแบ่งเนื้อหาออกเป็น 4 ตอน คือ ตอนที่ 1 สมรรถนะของผู้เรียนในศตวรรษที่ 21 ตอนที่ 2 การพัฒนาเครื่องมือวัด ตอนที่ 3 การสังเคราะห์งานวิจัย และตอนที่ 4 กรอบแนวคิดในการวิจัย มีรายละเอียดดังนี้

ตอนที่ 1 สมรรถนะของผู้เรียนในศตวรรษที่ 21

1.1 สมรรถนะของผู้เรียนในศตวรรษที่ 21

จากการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องในประเด็นสมรรถนะของผู้เรียนในศตวรรษที่ 21 ในเบื้องต้น พบว่า ประเทศต่าง ๆ ได้แก่ ไทย สหรัฐอเมริกา สิงคโปร์ ฮองกง ญี่ปุ่น และเกาหลีใต้ ได้กำหนดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ไว้ดังนี้

สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของไทย

แผนการศึกษาแห่งชาติ พ.ศ. 2560-2579 (สำนักงานเลขาธิการสภาการศึกษา, 2560) ได้กำหนดวิสัยทัศน์ไว้ว่า “คนไทยทุกคนได้รับการศึกษาและเรียนรู้ตลอดชีวิตอย่างมีคุณภาพ ดำรงชีวิตอย่างเป็นสุข สอดคล้องกับหลักปรัชญาของเศรษฐกิจพอเพียง และการเปลี่ยนแปลงของโลกศตวรรษที่ 21” โดยมีวัตถุประสงค์ในการจัดการศึกษา 4 ประการ คือ 1) เพื่อพัฒนาระบบและกระบวนการ จัดการศึกษาที่มีคุณภาพและมีประสิทธิภาพ 2) เพื่อพัฒนาคนไทยให้เป็นพลเมืองดี มีคุณลักษณะ ทักษะ และสมรรถนะที่สอดคล้องกับบทบัญญัติของรัฐธรรมนูญแห่งราชอาณาจักรไทย พระราชบัญญัติการศึกษาแห่งชาติและยุทธศาสตร์ชาติ 3) เพื่อพัฒนาสังคมไทยให้เป็นสังคมแห่งการเรียนรู้ และคุณธรรม จริยธรรม รู้รักสามัคคี และร่วมมือผนึกกำลังมุ่งสู่การพัฒนาประเทศอย่างยั่งยืน ตามหลักปรัชญาของเศรษฐกิจพอเพียง และ 4) เพื่อนำประเทศไทยก้าวข้ามกับดักประเทศที่มีรายได้ปานกลาง และความเหลื่อมล้ำภายในประเทศลดลง อย่างไรก็ตาม เพื่อให้การดำเนินบรรลุวิสัยทัศน์และจุดมุ่งหมายในการจัดการศึกษาดังกล่าวข้างต้น แผนการศึกษาแห่งชาติได้วางเป้าหมายไว้ 2 ด้าน คือ เป้าหมายด้านผู้เรียน (learner aspirations) และเป้าหมายของการจัดการศึกษา (aspirations)

อย่างไรก็ดีการกำหนดเป้าหมายด้านผู้เรียน ในแผนการศึกษาแห่งชาติ มุ่งพัฒนาผู้เรียนทุกคนให้มีคุณลักษณะและทักษะการเรียนรู้ ในศตวรรษที่ 21 ประกอบด้วยทักษะที่เรียกตามคำย่อว่า 3Rs + 8Cs ประกอบด้วย ทักษะและคุณลักษณะต่อไปนี้

3Rs ได้แก่ อ่านออก (Reading) เขียนได้ (WRiting) และ คิดเลขเป็น (ARithmetics)

8Cs ประกอบด้วย

- 1) ทักษะด้านการคิดอย่างมีวิจารณญาณ และทักษะในการแก้ปัญหา (critical thinking and problem solving)
- 2) ทักษะด้านการสร้างสรรค์และนวัตกรรม (creativity and innovation)
- 3) ทักษะด้านความเข้าใจต่างวัฒนธรรม ต่างกระบวนทัศน์ (cross-cultural understanding)

- 4) ทักษะด้านการทำงานร่วมกัน การทำงานเป็นทีม และภาวะผู้นำ (collaboration, teamwork and leadership)
- 5) ทักษะด้านการสื่อสาร สารสนเทศ และการรู้เท่าทันสื่อ (communications, information and media literacy)
- 6) ทักษะด้านคอมพิวเตอร์และ เทคโนโลยีสารสนเทศและการสื่อสาร (computing and ICT literacy)
- 7) ทักษะอาชีพ และทักษะ การเรียนรู้ (career and learning skills)
- 8) ความมีเมตตา กรุณา มีวินัย คุณธรรม จริยธรรม (compassion)

สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของสหรัฐอเมริกา

Partnership for 21st Century Learning หรือ P21 (2019) ได้พัฒนารอบแนวคิดการเรียนรู้ในศตวรรษที่ 21 โดยครู ผู้เชี่ยวชาญทางการศึกษา และผู้นำทางภาคธุรกิจในประเทศสหรัฐอเมริกา เพื่อที่กำหนดและแสดงให้เห็นถึงความรู้และทักษะที่ผู้เรียนจำเป็นต้องมีเพื่อความสำเร็จในอาชีพและการทำงาน ด้วยเหตุนี้จากความร่วมมือของกลุ่มบุคคลและองค์กรดังกล่าวข้างต้น จึงได้กำหนดกรอบแนวคิดการเรียนรู้ในศตวรรษที่ 21 ซึ่งประกอบด้วยผลการเรียนรู้ของผู้เรียนในศตวรรษที่ 21 มีรายละเอียดดังนี้

วิชาแกน เป็นเนื้อหาหลักสำคัญที่ผู้เรียนทุกคนจำเป็นต้องรู้ ประกอบด้วย

1. ภาษาอังกฤษ การอ่าน หรือศิลปะการใช้ภาษา (english, reading, or language arts)
2. ภาษาสำคัญของโลก (world languages)
3. ศิลปะ (arts)
4. คณิตศาสตร์ (mathematics)
5. เศรษฐศาสตร์ (economics)
6. วิทยาศาสตร์ (science)
7. ภูมิศาสตร์ (geography)
8. ประวัติศาสตร์ (history)
9. การปกครองและหน้าที่พลเมือง (government and civics)

แนวคิดสำคัญในศตวรรษที่ 21 นอกเหนือจากวิชาแกนที่ผู้เรียนจำเป็นต้องรู้แล้ว ควรมีการส่งเสริมความเข้าใจเกี่ยวกับเนื้อหาเชิงวิชาการในระดับที่สูงขึ้น ในลักษณะของสหวิทยาการ (interdisciplinary) ที่บูรณาการเข้าไปกับวิชาแกน ประกอบด้วย

1. จิตสำนึกต่อโลก (global awareness)
 - 1.1 ใช้ทักษะในศตวรรษที่ 21 เพื่อสร้างความเข้าใจและแก้ปัญหาในประเด็นที่เกี่ยวข้องกับโลก
 - 1.2 เรียนรู้การทำงานร่วมกันกับบุคคลอื่นที่มีความหลากหลายทางวัฒนธรรม ศาสนา และวิถีชีวิต มีการพูดคุยสื่อสารทั้งในระดับบุคคล การทำงาน และชุมชน
 - 1.3 เข้าใจผู้ที่มีวัฒนธรรมและเชื้อชาติอื่น รวมถึงผู้ที่ไม่ได้ใช้ภาษาอังกฤษ
2. ความรู้พื้นฐานด้านการเงิน เศรษฐกิจ ธุรกิจ และการเป็นผู้ประกอบการ (financial, economic, business, and entrepreneurial literacy)
 - 2.1 รู้ถึงตัวเลือกที่เหมาะสมต่อสภาพทางเศรษฐกิจของตนเอง
 - 2.2 เข้าใจบทบาทของเศรษฐกิจในสังคม

- 2.3 ใช้ทักษะผู้ประกอบการเพื่อเพิ่มประสิทธิภาพการทำงานและตัวเลือกอาชีพ
3. ความรู้พื้นฐานด้านพลเมือง (civic literacy)
- 3.1 มีส่วนร่วมในการเป็นพลเมืองด้วยการรู้วิธีการต่อการรับข้อมูลข่าวสาร และเข้าใจกระบวนการทางราชการ
- 3.2 ปฏิบัติตนตามสิทธิและหน้าที่ของความเป็นพลเมืองในระดับท้องถิ่น ระดับจังหวัด ระดับชาติ และระดับโลก
- 3.3 เข้าใจความเกี่ยวข้องกันระหว่างการตัดสินใจของพลเมืองในระดับท้องถิ่นและระดับโลก
4. ความรู้พื้นฐานด้านสุขภาพ (health literacy)
- 4.1 การได้รับ การแปลความหมาย และเข้าใจเกี่ยวกับข้อมูลสารสนเทศและการบริหารเกี่ยวกับพื้นฐานทางสุขภาพ รวมถึงการใช้ข้อมูลและบริการในหลากหลายวิธีการเพื่อส่งเสริมสุขภาพ
- 4.2 เข้าใจการตรวจวัดเพื่อป้องกันสุขภาพทางกายและทางใจ การรับประทานอาหารที่เหมาะสม มีคุณค่า การหลีกเลี่ยงความเสี่ยง และลดความเครียด
- 4.3 ใช้ข้อมูลสารสนเทศเพื่อตัดสินใจเกี่ยวกับเรื่องสุขภาพได้อย่างเหมาะสม
- 4.4 กำหนดและติดตามเป้าหมายสุขภาพของตนเองและครอบครัว
- 4.5 เข้าใจประเด็นความปลอดภัยและปัญหาสุขภาพในระดับชาติและนานาชาติ
5. ความรู้พื้นฐานด้านสิ่งแวดล้อม (environmental literacy)
- 5.1 แสดงให้เห็นถึงความรู้และความเข้าใจเกี่ยวกับสิ่งแวดล้อม เหตุการณ์ และเงื่อนไขที่มีผลต่ออากาศ บรรยากาศ พื้นดิน อาหาร หลังกาน น้ำ และระบบนิเวศน์
- 5.2 แสดงให้เห็นถึงความรู้และความเข้าใจเกี่ยวกับผลกระทบทางสังคมที่มีต่อธรรมชาติของโลก เช่น การเกิดของประชากร การพัฒนาประชากร อัตราการใช้ทรัพยากร เป็นต้น
- 5.3 ตรวจสอบและวิเคราะห์ประเด็นทางสิ่งแวดล้อม และสร้างข้อสรุปเกี่ยวกับการแก้ปัญหาที่มีประสิทธิผล
- 5.4 การมีส่วนร่วมในการแก้ไขปัญหาสิ่งแวดล้อม

ทักษะการเรียนรู้และนวัตกรรม (learning and innovation skills)

ทักษะการเรียนรู้และนวัตกรรม เป็นทักษะที่ได้รับการยอมรับว่าเป็นสิ่งที่เตรียมให้ผู้เรียนมีความพร้อมในการใช้ชีวิตและการทำงานในสภาพแวดล้อมที่มีความซับซ้อนของศตวรรษที่ 21 โดยมุ่งไปที่การเตรียมผู้เรียนให้มีความคิดสร้างสรรค์ การคิดอย่างมีวิจารณญาณ การสื่อสาร และความร่วมมือ มีรายละเอียดดังนี้

1. ความคิดสร้างสรรค์และนวัตกรรม (creativity and innovation) ประกอบด้วย

1.1 การคิดอย่างสร้างสรรค์ (think creatively)

- 1) การใช้เทคนิคในการสร้างความคิดที่หลากหลาย เช่น การระดมสมอง
- 2) สร้างความคิดที่แปลกใหม่โดย โดยอาจเป็นการปรับปรุงจากแนวคิดเดิม หรือแนวคิดที่ต่างจากเดิมอย่างสิ้นเชิง
- 3) ปรับปรุง วิเคราะห์ และประเมินผลการคิดของตนเอง เพื่อพัฒนาไปสู่การคิดอย่างสร้างสรรค์

1.2 การทำงานอย่างสร้างสรรค์ร่วมกับผู้อื่น (work creatively with others)

- 1) พัฒนา นำไปใช้ และสื่อสารความคิดใหม่กับผู้อื่นได้อย่างมีประสิทธิภาพ

2) เปิดใจและรับฟังความคิดเห็นมุมมองใหม่ โดยการให้ข้อมูลและข้อมูลย้อนกลับในการทำงาน

3) แสดงให้เห็นถึงความคิดริเริ่มและความคิดสร้างสรรค์ในการทำงาน และทำความเข้าใจกับข้อจำกัดในโลกแห่งความเป็นจริงเมื่อนำแนวคิดใหม่มาใช้

4) มองความล้มเหลวเป็นโอกาสในการเรียนรู้ เข้าใจว่าการสร้างสรรค์และนวัตกรรมเป็นสิ่งที่ต้องใช้เวลา กระบวนการที่เกิดความสำเร็จและความผิดพลาดเล็ก ๆ จะนำไปสู่การสร้างสรรค์นวัตกรรมได้

1.3 การนำนวัตกรรมไปใช้ (implement innovations)

ลงมือทำตามความคิดสร้างสรรค์ เพื่อสร้างผลงานนวัตกรรมที่เป็นรูปธรรม และประโยชน์

2. การคิดอย่างมีวิचारณญาณและการแก้ไขปัญหา (critical thinking and problem solving)

ประกอบด้วย

2.1 การให้เหตุผลอย่างมีประสิทธิภาพ (reason effectively)

- 1) ใช้เหตุผลหลากหลายประเภทที่เหมาะสมในการแก้ไขปัญหา
- 2) ใช้การคิดที่เป็นระบบ
- 3) วิเคราะห์ส่วนต่าง ๆ ที่มีปฏิสัมพันธ์กับส่วนอื่น ก่อนให้เกิดผลลัพธ์ที่ซับซ้อนได้อย่างไร
- 4) การตัดสินใจและการตัดสินใจ
- 5) การวิเคราะห์และประเมินเหตุการณ์ ข้อโต้แย้ง การกล่าวอ้าง และความเชื่อได้อย่างมีประสิทธิภาพ

ประกอบด้วย

- 6) การวิเคราะห์และประเมินทางเลือกที่สำคัญ
- 7) การสังเคราะห์และสร้างความเชื่อมโยงระหว่างข้อมูลสารสนเทศและข้อโต้แย้ง
- 8) การแปลความหมายของข้อมูลสารสนเทศและสรุปข้อมูลจากฐานการวิเคราะห์ที่ดีที่สุด
- 9) การสะท้อนกระบวนการและประสบการณ์เรียนรู้อย่างมีประสิทธิภาพ

2.2 การแก้ไขปัญหา (solve problems)

- 1) การแก้ไขปัญหาที่มีความหลากหลายทั้งในรูปแบบดั้งเดิมและที่เป็นนวัตกรรม
- 2) ระบุและถามคำถามที่มีนัยสำคัญได้อย่างชัดเจน และนำมาซึ่งวิธีการแก้ปัญหาคือดีที่สุด

3. การสื่อสารและการร่วมมือทำงาน (communication and collaboration) ประกอบด้วย

3.1 การสื่อสารที่ชัดเจน (communicate clearly)

1) สื่อสารความคิดได้อย่างมีประสิทธิภาพด้วยวิธีการพูด การเขียน ทักษะการสื่อสารด้วยอวัจนภาษาในรูปแบบและบริบทที่หลากหลาย

2) การฟังเพื่อจับใจความได้อย่างมีประสิทธิภาพ ซึ่งรวมถึงความรู้ คำนิยาม ทศนคติ และเจตนา

- 3) ใช้การสื่อสารเพื่อวัตถุประสงค์ที่หลากหลาย
- 4) ใช้สื่อและเทคโนโลยีที่หลากหลาย และรู้วิธีการตัดสินใจ และประเมินผลกระทบ
- 5) สื่อสารอย่างมีประสิทธิภาพในสภาพแวดล้อมที่หลากหลาย
- 6) ให้ความร่วมมือกับผู้อื่น
- 7) แสดงให้เห็นถึงความสามารถในการทำงานอย่างมีประสิทธิภาพและเคารพสมาชิกในกลุ่มที่มีความหลากหลาย

7) มีความยืดหยุ่นและความเต็มใจในการที่จะช่วยเหลือ ประนีประนอมตามความจำเป็น เพื่อบรรลุเป้าหมายร่วมกัน

8) แบ่งความรับผิดชอบในการทำงานของกลุ่มร่วมกัน และให้ความสำคัญต่อสมาชิกทุกคนในทีม

ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี (information, media, and technology skills)

บุคคลในศตวรรษที่ 21 อยู่ในสภาพแวดล้อมที่ขับเคลื่อนด้วยสื่อและเทคโนโลยี ที่ถูกสร้างด้วยคุณลักษณะที่หลากหลาย รวมถึง 1) การเข้าถึงสารสนเทศจำนวนมาก 2) การเปลี่ยนแปลงที่รวดเร็วของเครื่องมือทางเทคโนโลยี และ 3) ความสามารถในการทำงานร่วมกันและมีส่วนร่วมที่กับบุคคลที่ไม่เคยร่วมมาก่อน พลเมืองและแรงงานที่มีประสิทธิภาพของศตวรรษที่ 21 จำเป็นต้องสามารถแสดงทักษะการทำงาน การคิดวิเคราะห์ที่เกี่ยวข้องกับสารสนเทศ สื่อ และเทคโนโลยี มีรายละเอียดดังนี้

1. การรู้เท่าทันสารสนเทศ (information literacy) ประกอบด้วย

1.1 การเข้าถึงและประเมินสารสนเทศ (access and evaluate information)

- 1) การเข้าถึงสารสนเทศได้อย่างมีประสิทธิภาพและประสิทธิผล
- 2) การประเมินสารสนเทศอย่างมีวิจารณญาณและเต็มความสามารถ

1.2 การใช้และการจัดการสารสนเทศ (use and manage information)

- 1) ใช้สารสนเทศได้อย่างเหมาะสมและสร้างสรรค์สำหรับประเด็นหรือปัญหาที่กำลังเผชิญ
- 2) การจัดการกับข้อมูลที่มาจากแหล่งข้อมูลที่หลากหลาย
- 3) การเข้าถึงและใช้สารสนเทศโดยตั้งอยู่บนพื้นฐานความเข้าใจเกี่ยวกับประเด็นจริยธรรมและ

กฎหมาย

2. การรู้เท่าทันสื่อ (media literacy) ประกอบด้วย

2.1 การวิเคราะห์สื่อ (analyze media)

- 1) มีความเข้าใจว่าข้อความจากสื่อถูกสร้างอย่างไร เหตุผลใด และเพื่อวัตถุประสงค์ใด
- 2) ตรวจสอบว่าคุณแต่ละบุคคลมีการตีความข้อความจากสื่อ ค่านิยม แตกต่างกันอย่างไร มุมมองที่ได้บรรจุหรือตัดออกไปอย่างไร รวมถึงสื่อนั้นมีอิทธิพลต่อความเชื่อและพฤติกรรมอย่างไร
- 3) การเข้าถึงและใช้สื่อโดยตั้งอยู่บนพื้นฐานความเข้าใจเกี่ยวกับประเด็นจริยธรรมและ

กฎหมาย

2.2 การสร้างผลงานสื่อ (create media products)

- 1) มีความเข้าใจและใช้ประโยชน์จากเครื่องมือในการสร้างสื่อได้อย่างเหมาะสมมากที่สุดตามคุณลักษณะของเครื่องมือ
- 2) มีความเข้าใจและใช้ประโยชน์สื่อในการนำเสนอและตีความสื่อที่มีความหลากหลายของสภาพแวดล้อมและพฤติกรรมให้มีความเหมาะสมมากที่สุด
- 3) การรู้เท่าทันเทคโนโลยีสารสนเทศ (ICT literacy)

2.3 การใช้เทคโนโลยีได้อย่างมีประสิทธิภาพ (apply technology effectively)

- 1) การใช้เทคโนโลยีในฐานะที่เป็นเครื่องมือสำหรับการวิจัย จัดการ ประเมิน และสารสนเทศเพื่อการสื่อสาร

2) การใช้เทคโนโลยีดิจิทัลต่าง ๆ เพื่อการสื่อสาร เครื่องมือในการเชื่อมต่อเครือข่าย และสังคมเครือข่ายได้อย่างเหมาะสม เพื่อการเข้าถึง จัดการ บูรณาการ ประเมิน และสร้างสารสนเทศ เพื่อทำหน้าที่ในเศรษฐกิจฐานความรู้

3) การเข้าถึงและใช้เทคโนโลยีสารสนเทศโดยตั้งอยู่บนพื้นฐานความเข้าใจเกี่ยวประเด็นจริยธรรมและกฎหมาย

ทักษะชีวิตและการทำงาน (life and career skills)

สภาพแวดล้อมของชีวิตและการทำงานในปัจจุบันนอกจากความในเนื้อหาและทักษะการคิดแล้ว ผู้เรียนยังต้องมีความสามารถในการนำพาในสภาพแวดล้อมการทำงานและอาชีพที่มีความซับซ้อน ต้องใช้สารสนเทศในการแข่งขัน ผู้เรียนจึงจำเป็นต้องได้รับการพัฒนาทักษะชีวิตและการทำงาน มีรายละเอียดดังนี้

1. ความยืดหยุ่นและความสามารถในการปรับตัว (flexibility and adaptability)

1.1 การปรับตัวต่อการเปลี่ยนแปลง (adapt to change)

- 1) การปรับตัวต่อบทบาทที่หลากหลาย ความรับผิดชอบต่องาน ตารางงาน และบริบท
- 2) การทำงานอย่างมีประสิทธิภาพในสภาพที่มีความคลุมเครือ และมีการเปลี่ยนแปลงลำดับ

ความสำคัญ

1.2 มีความยืดหยุ่น (be flexible)

- 1) ร่วมให้ข้อมูลย้อนกลับที่มีประสิทธิภาพ
- 2) การจัดการในเชิงบวก เมื่อได้รับคำวิจารณ์ และความล้มเหลว
- 3) เข้าใจ ต่อรอง และสร้างสมดุลระหว่างมุมมองและความเชื่อที่หลากหลาย เพื่อให้สามารถแก้ไขปัญหาได้โดยเฉพาะในสภาพแวดล้อมที่มีความหลากหลายทางวัฒนธรรม

2. ความคิดริเริ่มและการชี้นำตนเอง (initiative and self-direction)

2.1 การกำหนดเป้าหมายและเวลา (manage goals and time)

- 1) การกำหนดเป้าหมาย เกณฑ์ความสำเร็จที่เป็นรูปธรรมและนามธรรม
- 2) สร้างสมดุลระหว่างเป้าหมายทางยุทธวิธี (ระยะสั้น) และเป้าหมายเชิงกลยุทธ์ (ระยะยาว)
- 3) ใช้เวลาและจัดการปริมาณงานได้อย่างมีประสิทธิภาพ

2.2 การทำงานอย่างอิสระ (work independently)

- 1) การกำกับติดตาม การกำหนด จัดลำดับความสำคัญ และการทำงานให้สำเร็จโดยไม่ต้องมีการกำกับดูแล
- 2) เรียนรู้แบบนำตนเอง
- 3) มีทักษะความเชี่ยวชาญพื้นฐาน และ/หรือตามหลักสูตรเพื่อสำรวจและขยายการเรียนรู้ของบุคคลใด และมีโอกาสในการเรียนรู้ในระดับเชี่ยวชาญ
- 4) แสดงให้เห็นถึงความริเริ่มในการพัฒนาทักษะไปสู่ระดับเชี่ยวชาญ
- 5) แสดงให้เห็นถึงความมุ่งมั่นว่าการเรียนรู้เป็นกระบวนการที่เกิดขึ้นตลอดชีวิต
- 6) สะท้อนประสบการณ์ที่ผ่านมาอย่างใคร่ครวญเพื่อนำไปสู่ความก้าวหน้าในอนาคต

3. ทักษะทางสังคมและการเรียนรู้ข้ามวัฒนธรรม (social and cross-cultural skills)

3.1 การมีปฏิสัมพันธ์กับผู้อื่นอย่างมีประสิทธิภาพ (interact effectively with others)

- 1) รู้ว่าเวลาใดควรฟังเวลาใดควรพูด
- 2) ปฏิบัติตนให้น่านับถือและมีความเป็นมืออาชีพ

3.2 การทำงานในทีมที่หลากหลายได้อย่างมีประสิทธิภาพ (work effectively in diverse teams)

- 1) เคารพวัฒนธรรมที่แตกต่าง และทำงานกับบุคคลอื่นที่มีภูมิหลังทางสังคมและวัฒนธรรมที่แตกต่างได้อย่างมีประสิทธิภาพ
- 2) เปิดใจในการรับความคิดและค่านิยมที่แตกต่าง
- 3) ใช้ประโยชน์จากความแตกต่างทางสังคมและวัฒนธรรมเพื่อสร้างความคิดใหม่ ๆ ทั้งนวัตกรรมและคุณภาพของงาน

4. การเพิ่มผลผลิตและความรับผิดชอบที่ตรวจสอบได้ (productivity and accountability)

4.1 การบริหารโครงการ (manage projects)

- 1) กำหนดและบรรลุเป้าหมาย แม้ว่าจะพบกับอุปสรรคและแรงกดดัน
- 2) การจัดลำดับความสำคัญ การวางแผน และการจัดการให้งานบรรลุตามผลที่ต้องการ

4.2 ผลผลิต (produce results)

แสดงให้เห็นถึงคุณลักษณะเพิ่มเติมที่เกี่ยวข้องกับการผลิตผลงานที่มีคุณภาพสูง โดยรวมถึงความสามารถ ดังนี้

- 1) ทำงานในเชิงบวกและมีจริยธรรม
- 2) การจัดการเวลาและโครงการอย่างมีประสิทธิภาพ
- 3) การทำงานได้อย่างหลากหลาย
- 4) การมีส่วนร่วมอย่างกระตือรือร้น มีความน่าเชื่อถือ และตรงเวลา
- 5) แสดงความเป็นมืออาชีพ และการแสดงออกอย่างเหมาะสม
- 6) ทำงานร่วมกันและให้ความร่วมมือกับทีมอย่างมีประสิทธิภาพ
- 7) เคารพและเห็นคุณค่าความหลากหลายของทีม
- 8) แสดงความรับผิดชอบที่ตรวจสอบได้ต่อผลที่เกิดขึ้น

5. ความเป็นผู้นำและความรับผิดชอบ (leadership and responsibility)

5.1 การแนะนำและนำผู้อื่น (guide and lead others)

1) ใช้การสื่อสารระหว่างบุคคลและทักษะการแก้ปัญหาเพื่อแนะนำผู้อื่นไปสู่เป้าหมายที่กำหนดไว้

- 2) เพิ่มจุดแข็งของผู้อื่นให้บรรลุเป้าหมายร่วมกัน
- 3) สร้างแรงบันดาลใจให้ผู้อื่นด้วยการให้ตัวอย่างที่ดีและเห็นแก่ส่วนรวม
- 4) แสดงให้เห็นถึงการยึดมั่นและมีจริยธรรมในการใช้อิทธิพลหรืออำนาจ

5.2 การแสดงความรับผิดชอบต่อบุคคลอื่น (be responsible to others)

การปฏิบัติด้วยความรับผิดชอบต่อความสนใจของชุมชนขนาดใหญ่

สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของสิงคโปร์

กระทรวงศึกษาธิการของประเทศสิงคโปร์ ได้กำหนด สมรรถนะของผู้เรียนในศตวรรษที่ 21 เพื่อเตรียมพร้อมผู้เรียนในอนาคต โดยอาศัยความร่วมมือระหว่างโรงเรียนและผู้ปกครอง ความรู้และทักษะต้องได้รับการสนับสนุนจากคุณค่า (values) คุณค่าเป็นสิ่งที่กำหนดคุณลักษณะของบุคคล จึงได้มีการกำหนด ความเชื่อ ทศนคติ และการปฏิบัติของบุคคล เป็นกรอบแนวคิดหลักของสมรรถนะของผู้เรียนในศตวรรษที่ 21 (Ministry of Education, Singapore., 2018) ดังภาพที่ 2.1



ภาพที่ 2.1 กรอบแนวคิดหลักของสมรรถนะของผู้เรียนในศตวรรษที่ 21 ของสิงคโปร์
ที่มา: Ministry of Education, Singapore. (2018)

ค่านิยมที่เป็นแกนหลักของสมรรถนะในศตวรรษที่ 21 (values at the core of 21st century competencies) ประกอบด้วย 6 ค่านิยมหลัก คือ

1. ความเคารพ (respect) ผู้เรียนเป็นบุคคลที่มีความเคารพ เมื่อผู้เรียนมีความเชื่อในค่านิยมของตนเอง และค่านิยมที่อยู่ภายในของบุคคลอื่น
2. ความรับผิดชอบ (responsibility) ผู้เรียนเป็นบุคคลที่มีความรับผิดชอบ เมื่อมีความตระหนักรู้ต่อหน้าที่ของตน ครอบครัว สังคม ประเทศชาติ และโลก โดยการเติมเต็มความรับผิดชอบด้วยความรักและความมุ่งมั่น
3. ความซื่อสัตย์ (integrity) ผู้เรียนเป็นบุคคลที่มีความซื่อสัตย์ หากยึดมั่นในหลักจริยธรรมและมีคุณธรรม กล่าวที่จะยืนหยัดเพื่อสิ่งที่ถูกต้อง
4. ความห่วงใย (care) ผู้เรียนเป็นบุคคลที่มีความห่วงใย ถ้าผู้เรียนแสดงถึงความเมตตา ความเห็นอกเห็นใจ และมีส่วนช่วยให้ชุมชนและโลกดีขึ้น
5. ความยืดหยุ่น (resilience) ผู้เรียนเป็นบุคคลที่มีความยืดหยุ่น ถ้ามีความมั่นคงทางอารมณ์และมีความมุ่งมั่นต่อสิ่งที่ท้าทายที่กำลังเผชิญอยู่ โดยแสดงให้เห็นถึงความกล้าหาญ การมองโลกในแง่ดี ความสามารถในการปรับตัว และไหวพริบในการแก้ปัญหา
6. ความสามัคคี (harmony) ผู้เรียนเป็นบุคคลที่เห็นค่านิยมของความสามัคคี หากสามารถหาความสุขภายใน ส่งเสริมการทำงานร่วมกันในสังคม เห็นค่านิยมของความเป็นหนึ่งเดียว และความหลากหลายทางสังคมวัฒนธรรม

สมรรถนะทางสังคมและอารมณ์ (social and emotional competencies)

เป็นส่วนของวงแหวนตรงกลางภาพเป็นทักษะที่มีความสำคัญสำหรับผู้เรียนเพื่อตระหนักและจัดการอารมณ์ของตนเอง พัฒนาความห่วงใยและใส่ใจผู้อื่น สร้างความรับผิดชอบต่อการตัดสินใจ สร้างความสัมพันธ์เชิงบวก รวมถึงการจัดการกับสถานการณ์ที่ท้าทายได้อย่างมีประสิทธิภาพ มีรายละเอียดดังนี้

1. ความตระหนักต่อตนเอง (self-awareness) ผู้เรียนเป็นบุคคลที่มีความตระหนักต่อตนเอง หากมีความเข้าใจในอารมณ์ จุดแข็ง จุดอ่อน และความชอบของตนเอง

2. การจัดการตนเอง (self-management) ผู้เรียนเป็นบุคคลที่จัดการตนเองได้อย่างมีประสิทธิภาพ หากเป็นผู้ที่มีความสามารถในการบริหารจัดการอารมณ์ของตนเองได้ ด้วยการสร้างแรงจูงใจ

3. ความตระหนักรู้ทางสังคม (social awareness) ผู้เรียนเป็นบุคคลที่มีความตระหนักรู้ทางสังคม หากผู้เรียนมีความสามารถในการมองเห็นมุมมองที่แตกต่าง ยอมรับและเห็นคุณค่าในความหลากหลาย เอาใจใส่และเคารพผู้อื่น

4. การบริหารความสัมพันธ์ (relationship management) ผู้เรียนเป็นบุคคลที่สามารถบริหารความสัมพันธ์ได้อย่างมีประสิทธิภาพ ถ้ามีความสามารถในการสร้างและรักษาความสัมพันธ์ที่เข้มแข็ง ผ่านการสื่อสารที่มีประสิทธิภาพ และสามารถทำงานร่วมกับผู้อื่นเพื่อแก้ไขปัญหาและให้ความช่วยเหลือ

5. การตัดสินใจบนความรับผิดชอบต่อ (responsible decision-making) ผู้เรียนสามารถตัดสินใจบนความรับผิดชอบต่อ หากมีความสามารถในการระบุและวิเคราะห์สถานการณ์ได้อย่างสมบูรณ์ โดยผู้เรียนสามารถสะท้อนการตัดสินใจที่ตั้งอยู่บนฐานของบุคคล คุณธรรม และจริยธรรม

สมรรถนะสำคัญในศตวรรษที่ 21 (emerging 21st century competencies)

ในขณะที่พื้นที่วงแหวนรอบนอกเป็นกรอบที่แสดงถึงสมรรถนะที่สำคัญในโลกยุคโลกาภิวัตน์ ประกอบด้วย

1. ความรู้การเป็นพลเมือง ความตระหนักต่อโลก และทักษะการเรียนรู้ข้ามวัฒนธรรม (civic literacy, global awareness & cross-cultural skills) ด้วยสภาพสังคมของสิงคโปร์มีความเป็นสากลทั้งในการใช้ชีวิตและการทำงานในต่างแดนมากยิ่งขึ้น ผู้เรียนจึงจำเป็นต้องมีมุมมองต่อโลกที่กว้างขึ้น และมีความสามารถในการทำงานกับผู้อื่นที่มีภูมิหลังบนความหลากหลายทางวัฒนธรรม ความแตกต่างทางความคิดและมุมมอง ในขณะเดียวกันผู้เรียนควรรู้ประเด็นปัญหาในระดับชาติ มีความภาคภูมิใจในความเป็นคนสิงคโปร์และมีความรับผิดชอบต่อชุมชน

2. การคิดอย่างมีวิจารณญาณและสร้างสรรค์ (critical and inventive thinking) เพื่อความพร้อมในอนาคต ผู้เรียนจำเป็นต้องมีความสามารถในการคิดอย่างมีวิจารณญาณ ประเมินความคิดเห็น และตัดสินใจ ผู้เรียนควรมีความต้องการในการเรียนรู้ สำรวจ และเตรียมการคิดนอกกรอบ ไม่กลัวต่อความผิดพลาด และกล้าเผชิญหน้าต่อสิ่งท้าทายที่พบเจอเป็นครั้งแรก

3. ทักษะการสื่อสาร การทำงานร่วมกัน และสารสนเทศ (communication, collaboration and information skills) การปฏิวัติของอินเทอร์เน็ต การได้รับข้อมูลสารสนเทศได้เพียงปลายนิ้ว ผู้เรียนจึงควรเรียนรู้ที่จะตั้งคำถาม รู้วิธีการกรองข้อมูล การสกัดข้อมูลที่เกี่ยวข้องและเป็นประโยชน์ ในขณะเดียวกันผู้เรียนจำเป็นต้องแยกแยะสิ่งที่สามารถทำให้เกิดอันตราย ปฏิบัติตามหลักจริยธรรมในโลกไซเบอร์ การทำงานในศตวรรษที่ 21 ต้องการคนที่สามารถทำงานร่วมกันโดยให้ความเคารพ มีความรับผิดชอบต่อการส่งต่อข้อมูล และตัดสินใจร่วมกันเพื่อให้บรรลุเป้าหมายของกลุ่ม โดยสามารถสื่อสารความคิดได้อย่างชัดเจนและมีประสิทธิภาพ

สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของเขตบริหารพิเศษฮ่องกง

เป้าหมายการเรียนรู้ของเขตบริหารพิเศษฮ่องกง คือการพัฒนาคุณภาพของคนในยุคใหม่ ให้สามารถอยู่รอด เป็นผู้นำ และเปลี่ยนแปลงสังคมที่มีความผันผวน ไม่แน่นอน ซับซ้อน และคลุมเครือ จึงมุ่งเน้นพัฒนาคุณภาพประชาชนใน 3 มิติ ดังนี้ (Cheng, Jackson & Lee, 2017)

1. ความรู้และทักษะ (knowledge and skills) เพื่อสร้างพื้นฐานสำหรับการทำงานและชีวิต ประกอบด้วยตัวอย่างเป้าหมายที่สำคัญคือ

- 1.1 ความสามารถในการใช้ภาษาที่สองและภาษาที่สาม
- 1.2 รู้ลึกและรู้กว้าง (T-shaped knowledge)
- 1.3 ความรู้ความเชี่ยวชาญในเทคโนโลยี
- 1.4 ความรู้เกี่ยวกับภูมิศาสตร์ เศรษฐกิจ ประวัติศาสตร์ และวัฒนธรรมของสังคม ประเทศชาติ

และโลก

2. คุณลักษณะ (attributes) เพื่อการดำรงอยู่ การใช้ประโยชน์ การประสบความสำเร็จ และการปฏิบัติ ตัวอย่างเป้าหมายที่สำคัญ เช่น

- 2.1 การคิดอย่างมีวิจารณญาณ
- 2.2 ความมีน้ำใจต่อกลุ่ม
- 2.3 การเห็นคุณค่าของความแตกต่าง
- 2.4 ความพยายาม
- 2.5 การสร้างสรรค์
- 2.6 ความอ่อนน้อมถ่อมตนและความมั่นใจในตนเอง
- 2.7 การจัดการตนเอง (self-management)
- 2.8 การควบคุมตนเอง (self-control)

3. ค่านิยม (values) การสร้างอนาคตที่ดีกว่า และนำไปสู่ชีวิตที่มีความหมาย

- 3.1 ความซื่อสัตย์
- 3.2 ความรับผิดชอบ
- 3.3 การรักษาสัญญา
- 3.4 การเคารพค่านิยมที่แตกต่าง
- 3.5 ความอดทนต่อความหลากหลาย
- 3.6 ความยุติธรรม
- 3.7 การปฏิบัติตามกฎหมาย
- 3.8 ความสงบสุข

สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของญี่ปุ่น

ได้กำหนดเป้าหมายเพื่อพัฒนาความสามารถทางปัญญา (intellectual) คุณธรรม (moral) และสมรรถภาพทางกาย (physical prowess) ดังนี้ (Cheng, 2017)

การรู้พื้นฐาน (basic literacy) ประกอบด้วย

ด้านภาษา (literacy) ด้านคำนวณ (numeracy) และการรู้เทคโนโลยีสารสนเทศ (information /communication technology literacy) รวมถึงมารยาทและจริยธรรมสำหรับการใช้เทคโนโลยีสารสนเทศ

ความสามารถทางการคิด (thinking ability) ประกอบด้วย

การสืบค้นและแก้ไขปัญหา (finding and solving problems) การสร้างสรรค์ (creativity) การคิดอย่างมีวิจารณญาณ (critical thinking) การคิดเชิงตรรกะ (logical thinking) อภิปัญญา (metacognition) และทักษะการเรียนรู้การปรับตัว (adaptive learning skills)

ความสามารถในการปฏิบัติตนต่อโลก (practical ability to act for the world) ประกอบด้วย

1) การปฏิบัติตนได้อย่างอิสระ (independence and autonomous action) ประกอบด้วย การเข้าใจตนเอง ความรับผิดชอบต่อตนเอง การส่งเสริมสุขภาพ ทักษะการตัดสินใจ และทักษะการวางแผนชีวิต

2) การสร้างความสัมพันธ์ (relationship building) ประกอบด้วย การให้ความร่วมมือและความรับผิดชอบ ความไว/การแสดงออก และการสร้างความสัมพันธ์กับผู้อื่นได้ดี

3) ความรับผิดชอบต่อการสร้างอนาคตที่ยั่งยืน (responsibility for building a sustainable future) ประกอบด้วย ความรับผิดชอบต่อ สิทธิ และการทำงาน ความเข้าใจต่อสังคม วัฒนธรรม และสิ่งแวดล้อม ตามธรรมชาติ การประยุกต์ใช้ภาษาและสารสนเทศ การประยุกต์ใช้ความรู้และเทคโนโลยี และการทักษะการค้นหาคำถามและแก้ไขปัญหา

สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของเกาหลีใต้ (Cheng, 2017)

1) สมรรถนะการจัดการตนเอง (self-management competency)

2) สมรรถนะความรู้/การประมวลผลสารสนเทศ (knowledge/information processing competency)

3) สมรรถนะการคิดสร้างสรรค์ (creative thinking competency)

4) สมรรถนะสุนทรียทางอารมณ์ (aesthetic-emotional competency)

5) ทักษะการสื่อสาร (communication skills)

6) สมรรถนะพลเมือง (civic competency)

นอกจากนี้ Greenstein (2012) ได้สังเคราะห์สมรรถนะในศตวรรษที่ 21 จากแนวคิดของนักการศึกษา สามารถแบ่งองค์ประกอบหลักสำคัญออกเป็น 3 องค์ประกอบ ดังนี้

ด้านการคิด (thinking) ประกอบด้วย 1) การคิดอย่างมีวิจารณญาณ (critical thinking) 2) การแก้ปัญหา (problem solving) 3) การสร้างสรรค์ (creating) และ 4) อภิปัญญา (metacognition)

ด้านปฏิบัติ (acting) ประกอบด้วย 1) การสื่อสาร (communicating) 2) การทำงานร่วมกัน (collaborating) 3) การรู้เท่าทันดิจิทัล (digital literacy) 4) การรู้เท่าทันเทคโนโลยี (technology literacy) 5) ความสามารถในการยืดหยุ่นและการปรับตัว (flexibility and adaptability) และ 6) ความคิดริเริ่มและการนำตนเอง (initiative and self-direction)

ด้านการใช้ชีวิตในโลก (living in the world) ประกอบด้วย 1) หน้าที่พลเมืองและความเป็นพลเมือง (civic responsibility and citizenship) 2) ความเข้าใจต่อโลก (global understanding) 3) ภาวะผู้นำและความรับผิดชอบ (leadership and responsibility) และ 4) ความพร้อมทางการเรียนและวิชาชีพ (college and career readiness)

จากผลการศึกษาสมรรถนะในศตวรรษที่ 21 จากประเทศต่าง ๆ ข้างต้น รวมถึงผลการสังเคราะห์ของ Greenstein (2012) ผู้วิจัยได้ดำเนินการสังเคราะห์สมรรถนะของผู้เรียนในศตวรรษที่ 21 เพื่อพิจารณาว่าสมรรถนะใดเป็นสมรรถนะที่สำคัญในศตวรรษที่ 21 ที่แต่ละประเทศมุ่งเน้นให้เกิดขึ้นกับผู้เรียน ผลการสังเคราะห์ ผู้วิจัยได้จัดกลุ่มสมรรถนะออกเป็น สมรรถนะด้านทักษะและความรู้พื้นฐาน ด้านทักษะและความรู้ในเนื้อหาขั้นสูง ด้านทักษะการคิดขั้นสูง ด้านทักษะต่อตนเองและสังคม ด้านทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี และด้านค่านิยมหลัก ผลการสังเคราะห์ดังตารางที่ 2.1

ตารางที่ 2.1 ผลการสังเคราะห์สมรรถนะของผู้เรียนในศตวรรษที่ 21

สมรรถนะ	สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของแต่ละประเทศ							รวม
	ไทย	USA	สิงคโปร์	ฮ่องกง	ญี่ปุ่น	เกาหลีใต้	Greenstein (2012)	
ทักษะและความรู้พื้นฐาน								
ภาษา (language)		✓		✓	✓			3
อ่านออก (reading)	✓	✓						2
เขียนได้ (writing)	✓							1
คิดเลขเป็น (arithmetics)	✓	✓			✓			3
ภูมิศาสตร์ เศรษฐกิจ ประวัติศาสตร์ (geography economics history)		✓		✓				2
ศิลปะ (arts)		✓						1
วิทยาศาสตร์ (science)		✓						1
การปกครองและหน้าที่พลเมือง (government and civics)		✓						1
รู้ลึกและรู้กว้าง (t-shaped knowledge)				✓				1
ทักษะและความรู้ในเนื้อหาขั้นสูง								
จิตสำนึกต่อโลก (global awareness)		✓	✓		✓	✓	✓	5*
ความรู้เรื่องการเงิน เศรษฐกิจ ธุรกิจและการเป็นผู้ประกอบการ (financial, economic, business, and entrepreneurial literacy)		✓						1
ความรู้เรื่องพลเมือง (civic literacy)		✓	✓			✓	✓	4*
ความรู้เรื่องสุขภาพ (health literacy)		✓						1
ความรู้เรื่องสิ่งแวดล้อม (environmental literacy)		✓						1
ทักษะการคิดขั้นสูง								
การคิดอย่างมีวิจารณญาณ (critical thinking)	✓	✓	✓	✓	✓		✓	6*
การแก้ปัญหา (solving problems)	✓	✓			✓		✓	4*
การสร้างสรรค์และนวัตกรรม (creativity and innovation)	✓	✓	✓	✓	✓	✓	✓	7*
ความคิดริเริ่ม (initiative)		✓	✓				✓	3

สมรรถนะ	สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของแต่ละประเทศ							รวม
	ไทย	USA	สิงคโปร์	ฮ่องกง	ญี่ปุ่น	เกาหลีใต้	Greenstein (2012)	
การคิดเชิงตรรกะ (logical thinking)					✓			1
อภิปัญญา (metacognition)					✓		✓	2
ทักษะต่อตนเองและสังคม								
การสื่อสาร (communications literacy/skill)	✓	✓	✓		✓	✓	✓	6*
การนำตนเอง (self-direction)						✓	✓	2
การจัดการตนเอง (self-management)			✓	✓		✓		3
การควบคุมตนเอง (self-control)				✓				1
ความตระหนักรู้ต่อตนเอง (self-awareness)			✓					1
ความตระหนักรู้ทางสังคม (social awareness)			✓					1
ทักษะทางสังคมและการเรียนรู้ข้าม วัฒนธรรม (social and cross-cultural skills)	✓	✓	✓					3
ทักษะอาชีพ และทักษะการเรียนรู้ (career and learning skills)	✓	✓					✓	3
ทักษะการเรียนรู้การปรับตัว (adaptive learning skills)					✓	✓	✓	3
ความรับผิดชอบที่ตรวจสอบได้ (accountability)		✓	✓					2
ทักษะด้านการทำงานเป็นทีม ความ ร่วมมือ (collaboration, teamwork)	✓	✓	✓				✓	4*
การบริหารความสัมพันธ์ (relationship management)			✓					1
ภาวะผู้นำ (leadership)	✓	✓				✓	✓	4*
ความรับผิดชอบ (responsibility)		✓	✓	✓		✓	✓	5*
การตัดสินใจบนความรับผิดชอบ (responsible decision-making)			✓					1
ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี								
การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy)	✓	✓	✓		✓	✓		5*
การรู้เท่าทันดิจิทัล (digital literacy)						✓	✓	2
การรู้เท่าทันเทคโนโลยี (technology literacy)	✓	✓		✓	✓	✓	✓	6*
การรู้เท่าทันการใช้คอมพิวเตอร์ (computing literacy)	✓							1

สมรรถนะ	สมรรถนะของผู้เรียนในศตวรรษที่ 21 ของแต่ละประเทศ							รวม
	ไทย	USA	สิงคโปร์	ฮ่องกง	ญี่ปุ่น	เกาหลีใต้	Greenstein (2012)	
ค่านิยมหลัก								
ความมีเมตตากรุณา	✓							1
ความมีวินัย และการปฏิบัติตามกฎหมาย	✓			✓				2
คุณธรรม จริยธรรม	✓							1
ความเคารพ			✓	✓				2
ความซื่อสัตย์			✓	✓				2
ความห่วงใย			✓					1
ความยืดหยุ่น			✓					1
ความสามัคคี			✓					1
การรักษาสัญญา				✓				1
ความยุติธรรม								
ความอดทนต่อความหลากหลาย								
ความสงบสุข								
สมรรถนะสุนทรียทางอารมณ์ (aesthetic-emotional competency)						✓		1

เมื่อกำหนดเกณฑ์ว่าหากมีเอกสารที่ระบุสมรรถนะตรงกันตั้งแต่ 4 เรื่องขึ้นไป นับได้ว่าเป็นสมรรถนะสำคัญ จากตารางที่ 2.1 ผลการสังเคราะห์สมรรถนะของผู้เรียนในศตวรรษที่ 21 พบว่า สมรรถนะสำคัญของผู้เรียนในศตวรรษที่ 21 ประกอบด้วย

1. ทักษะและความรู้ในเนื้อหาขั้นสูง

1.1 จิตสำนึกต่อโลก (global awareness)

1.2 ความรู้เรื่องพลเมือง (civic literacy)

2. ทักษะการคิดขั้นสูง

2.1 การคิดอย่างมีวิจารณญาณ (critical thinking)

2.2 การแก้ปัญหา (solving problems)

2.3 การสร้างสรรค์และนวัตกรรม (creativity and innovation)

3. ทักษะต่อตนเองและสังคม

3.1 การสื่อสาร (communications literacy/skill)

3.2 ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (collaboration, teamwork)

3.3 ภาวะผู้นำ (leadership)

3.4 ความรับผิดชอบ (responsibility)

4. ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี

4.1 การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy)

4.2 การรู้เท่าทันเทคโนโลยี (technology literacy)

ผู้วิจัยจึงได้คัดเลือกประเด็นดังกล่าวข้างต้นมาใช้ในการสังเคราะห์งานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

1.2 งานวิจัยที่เกี่ยวกับสมรรถนะของผู้เรียนในศตวรรษที่ 21

น้ำทิพย์ งามอาภาวิชย์ (2556) ได้พัฒนาแบบวัดทักษะแห่งศตวรรษที่ 21 มีวัตถุประสงค์วิจัย 3 ประการคือ 1) เพื่อพัฒนาและตรวจสอบคุณภาพของแบบวัด ทักษะแห่งศตวรรษที่ 21 ตามการรับรู้ของผู้เรียนมัธยมศึกษาตอนต้น โดยประยุกต์ใช้แนวคิดการเข้าถึง คุณลักษณะที่มุ่งวัดของแบบสอบ (2) เพื่อสร้างเกณฑ์ปกติ (3) เพื่อประเมินทักษะแห่งศตวรรษที่ 21 ตามการรับรู้ของผู้เรียน กลุ่มตัวอย่างเป็นผู้เรียนระดับมัธยมศึกษาตอนต้น (มัธยมศึกษาปีที่ 1-3) สังกัด สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน จำนวน 1,860 คน ได้มาจากการสุ่มแบบหลายขั้นตอน เครื่องมือที่ใช้ในการ ได้แก่ แบบวัดทักษะแห่งศตวรรษที่ 21 ชนิดมาตราประมาณค่า และชนิดมาตราประมาณค่าเชิงพฤติกรรม โดยประยุกต์แนวคิดการเข้าถึงคุณลักษณะที่มุ่งวัดของแบบสอบ วิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูป SPSS for windows, โปรแกรม MULTILOG, โปรแกรม LISREL for windows และโปรแกรม microsoft office excel ผลการวิจัยพบว่า 1) แบบวัดทักษะแห่งศตวรรษที่ 21 ที่สร้างขึ้นมีโครงสร้าง 3 องค์ประกอบ ได้แก่ ทักษะการเรียนรู้และนวัตกรรม (10 ข้อ) ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี (13 ข้อ) และ ทักษะชีวิตและการทำงาน (19 ข้อ) รวมทั้งสิ้นจำนวน 42 ข้อ มีค่าความเที่ยงเท่ากับ 0.948 และแบบวัด มีความตรงเชิงโครงสร้าง ($\chi^2 = 44.56$, $df = 33$, $p = 0.086$, $GFI = 0.99$, $AGFI = 0.98$, $RMSEA = 0.021$) 2) คะแนนปกติวิสัยของทักษะแห่งศตวรรษที่ 21 อยู่ในช่วงเปอร์เซ็นต์ไทล์ที่ 0 ถึง 99 และมีช่วงคะแนนที่ตั้งตั้งแต่ 15 - 82 และ 3) ผลการประเมินทักษะแห่งศตวรรษที่ 21 พบว่า ผู้เรียน มีทักษะแห่งศตวรรษที่ 21 อยู่ในระดับสูง ค่อนข้างสูง ค่อนข้างต่ำ และต่ำ โดยแต่ละระดับคิดเป็นร้อยละ 24.46, 26.36, 24.18 และ 25.00 ตามลำดับ

พิชญา ตีมี (2559) ได้ดำเนินการวิจัยเพื่อพัฒนาแนวทางการประเมินทักษะการเรียนรู้และนวัตกรรมของผู้เรียน ตามแนวคิดการประเมินผู้เรียนในศตวรรษที่ 21 โดยมีวัตถุประสงค์เพื่อพัฒนา ทดลองใช้ และประเมินคุณภาพแนวทางการประเมินของผู้เรียนตามแนวคิดการประเมินผู้เรียนในศตวรรษที่ 21 โดยใช้รูปแบบการวิจัยและพัฒนา ที่มีการผสมวิธีทั้งเชิงปริมาณและเชิงคุณภาพ ทักษะการเรียนรู้และนวัตกรรมของผู้เรียน ในการวิจัยครั้งนี้ประกอบด้วย 3 ด้าน ได้แก่ การสร้างสรรค์และนวัตกรรม การคิดอย่างมีวิจารณญาณและการแก้ปัญหา การสื่อสารและการร่วมมือ กลุ่มตัวอย่างในการทดลองใช้แนวทางการประเมิน ได้แก่ ผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 5 โรงเรียนพิริยาลัยจังหวัดแพร่ ปี การศึกษา 2558 จำนวน 41 คน ครูจำนวน 3 คน โดยใช้เทคนิคการวิจัยเชิงปฏิบัติการ (action research) ตามวงจร PAOR จำนวน 2 วงรอบ ผลการวิจัยพบว่า 1) แนวทางการประเมินที่พัฒนาขึ้น ประกอบด้วย โครงสร้างความสัมพันธ์กันของรายละเอียด 6 องค์ประกอบ ได้แก่ วัตถุประสงค์การประเมิน องค์ประกอบของทักษะและตัวชี้วัด เทคนิควิธีการประเมิน เครื่องมือที่ใช้ในการประเมิน การตัดสินใจประเมิน และการรายงานผลการประเมิน 2) ผลการทดลองใช้แนวทาง ประเมินวงรอบที่ 1 ผู้เรียนมีคะแนนเฉลี่ยของทักษะการเรียนรู้และนวัตกรรม เท่ากับ 64.44 และวงรอบที่ 2 มีคะแนนเฉลี่ย 69.49 ในภาพรวมผู้เรียนมีระดับทักษะการเรียนรู้และนวัตกรรมในระดับดีมากที่สุด คิดเป็นร้อยละ 65.88 และระดับปานกลาง คิดเป็นร้อยละ 34.16 3) ผลการประเมินคุณภาพของแนวทางการประเมินทักษะการเรียนรู้และนวัตกรรมของผู้เรียนตามแนวคิดการประเมินผู้เรียนใน ศตวรรษที่ 21 มีคุณภาพในภาพรวมอยู่ในระดับมาก

ปวีณา มะแซ. (2561) ได้พัฒนา หาคคุณภาพ และสร้างเกณฑ์ปกติของแบบวัดทักษะการรู้เท่าทันสื่อ ในศตวรรษที่ 21 ของนักเรียนชั้นมัธยมศึกษาตอนต้นโดยประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุ

วิภาค กลุ่มตัวอย่างที่ใช้ในการวิจัยเป็นนักเรียนชั้นมัธยมศึกษาตอนต้นสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 15 จำนวน 1,000 คน โดยวิธีการสุ่มแบบแบ่งชั้น เครื่องมือที่ใช้ในการวิจัย เป็นแบบวัดทักษะการรู้เท่าทันสื่อในศตวรรษที่ 21 ที่ผู้วิจัยพัฒนาขึ้นมีรูปแบบเป็นข้อสอบสถานการณ์ชนิดปรนัย 4 ตัวเลือก ตรวจสอบคุณภาพโดยใช้ทฤษฎีการทดสอบแบบดั้งเดิม ทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค ได้ตรวจสอบความเป็นเอกมิติโดยใช้วิธีวิเคราะห์องค์ประกอบเชิงยืนยัน (CFA) ตรวจสอบพารามิเตอร์ของข้อสอบ ได้แก่ อำนาจจำแนก (α), ความยาก (β) และสารสนเทศของแบบวัดด้วย Grade response model (GRM) และสร้างเกณฑ์ปกติของแบบวัดทักษะการรู้เท่าทันสื่อในศตวรรษที่ 21 แบบคะแนนที่ปกติ ผลการวิจัยพบว่า 1) แบบวัดที่พัฒนาขึ้น มี 5 องค์ประกอบ ได้แก่ ทักษะการเข้าถึง ทักษะการวิเคราะห์ ทักษะการประเมิน ทักษะการสร้างสรรค์ และทักษะการมีส่วนร่วม จำนวน 42 ข้อ มีความตรงเชิงพินิจอยู่ระหว่าง 0.60 – 1.00 ข้อสอบทุกข้อสามารถจำแนกกลุ่มสูงกลุ่มต่ำได้อย่างมีนัยสำคัญทางสถิติ และมีค่าความเที่ยง เท่ากับ 0.87 2) ผลการวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่า แบบวัดมีความเป็นเอกมิติ ทั้งนี้ ผลการตรวจสอบคุณภาพของแบบวัดตามทฤษฎีการตอบสนองข้อสอบ พบว่า ค่าอำนาจจำแนกรายข้อ (α) อยู่ระหว่าง 0.06 – 2.01 ค่าความยาก (β) ของแต่ละรายการคำตอบมีค่าเรียงลำดับจากน้อยไปมากทุกข้อ และค่าสารสนเทศของแบบวัดมีความเที่ยงเท่ากับ 0.89 และ 3) เกณฑ์ปกติของแบบวัดคะแนนที่ปกติมีค่าอยู่ระหว่าง 17 ถึง 83

ปาริชาติ อังกาบ (2561) ได้สร้าง หาคุณภาพ และสร้างเกณฑ์ปกติของแบบวัดทักษะชีวิตและอาชีพตามแนวคิดทักษะในศตวรรษที่ 21 สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย กลุ่มตัวอย่างที่ใช้ในการวิจัย คือ นักเรียนระดับชั้นมัธยมศึกษาตอนปลาย สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 25 จำนวน 618 คน ได้จากการสุ่มแบบหลายชั้นตอน ผลการวิจัยพบว่า 1) แบบวัดที่พัฒนาขึ้นมี 5 องค์ประกอบ ได้แก่ ด้านความยืดหยุ่นและการปรับตัว, ด้านการริเริ่มและการนำตนเอง, ด้านทักษะสังคมและการข้ามวัฒนธรรม, ด้านการผลิตงานและความรับผิดชอบตรวจสอบได้ และด้านภาวะผู้นำและความรับผิดชอบ องค์ประกอบ 8 ข้อ รวมจำนวน 40 ข้อ 2) คุณภาพของแบบวัด มีค่าดัชนีความสอดคล้อง (IOC) ระหว่าง 0.60 ถึง 1.00 ค่าอำนาจจำแนก ระหว่าง 0.319 ถึง 0.748 ค่าความเที่ยงทั้งหมด เท่ากับ 0.910 ผลการวิเคราะห์องค์ประกอบเชิงยืนยัน พบว่า แบบวัดมีความตรงเชิงโครงสร้าง และ 3) เกณฑ์ปกติในรูปคะแนนที่ปกติอยู่ระหว่าง T24 ถึง T72

จากการสำรวจงานวิจัยที่เกี่ยวข้องกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ในเบื้องต้นพบว่า มีงานวิจัยที่เป็นไปตามขอบเขตที่ผู้วิจัยกำหนดจำนวน 92 เล่ม ซึ่งในจำนวนนี้เป็นการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 โดยตรง และส่วนหนึ่งเป็นการพัฒนาเครื่องมือวัดอื่น ๆ แต่มีองค์ประกอบของเครื่องมือวัดบางองค์ประกอบเป็นไปตามกรอบของสมรรถนะของผู้เรียนในศตวรรษที่ 21 โดยผู้วิจัยจะนำงานวิจัยดังกล่าวมาทำการสังเคราะห์เพื่อสรุปเป็นองค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษาต่อไป

ตอนที่ 2 การพัฒนาเครื่องมือวัด

การวัดพฤติกรรมหรือความรู้สึกทางสังคมศาสตร์นั้นจำเป็นต้องอาศัยเครื่องมือซึ่งเป็นสิ่งเร้าให้ผู้ตอบแสดงความรู้สึกหรือพฤติกรรมออกมา โดยเครื่องมือที่มีคุณภาพในการวัดจิตลักษณะและพฤติกรรมจะต้องสามารถวัดได้ตรงกับสิ่งที่ผู้วิจัยตั้งใจจะวัดและผลการวัดที่ได้ต้องเชื่อถือได้ โดยผู้วิจัยแบ่งการนำเสนอออกเป็น 4 หัวข้อ คือ 2.1 ประเภทของเครื่องมือวัด 2.2 ขั้นตอนในการพัฒนาเครื่องมือวัด 2.3 การตรวจสอบคุณภาพเครื่องมือ และ 2.4 การสร้างเกณฑ์ปกติ ซึ่งมีรายละเอียดดังนี้

2.1 ประเภทของเครื่องมือวัด

สำหรับประเภทของเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ผู้วิจัยได้ศึกษารูปแบบเครื่องมือที่งานวิจัยส่วนใหญ่พัฒนาขึ้น ดังนี้

1. แบบทดสอบ (test)

แบบทดสอบสามารถแบ่งประเภทตามลักษณะข้อสอบได้เป็น 2 ประเภท คือ แบบสอบประเภทเลือกคำตอบ (selection type) และแบบสอบประเภทเสนอคำตอบ (supply type) มีรายละเอียด ดังนี้ (ณัฐภรณ์ หลาวทอง, 2551; เยาวดี วิบูลย์ศรี; 2556; ศิริชัย กาญจนวาสี, 2556; สมนึก ภัททิยธนี, 2560; สำนักวิชาการและมาตรฐานการศึกษา, 2557; Gronlund & Waugh, 2009; Miller, Linn & Gronlund, 2009)

1.1 แบบสอบประเภทเลือกคำตอบ (selection type)

เป็นแบบสอบที่กำหนดคำตอบไว้ให้ผู้สอบทำการเลือกคำตอบ โดยผู้สอบจะต้องอ่าน คิด และเลือกคำตอบที่ถูกต้อง การตรวจข้อสอบสามารถทำได้สะดวก การให้คะแนนมีความเป็นปรนัย (objective) ข้อสอบประเภทนี้ ได้แก่ ข้อสอบถูกผิด (true-false) ข้อสอบจับคู่ (matching) และข้อสอบเลือกตอบ (multiple choice) มีรายละเอียดดังนี้

1) ข้อสอบถูกผิด (true-false)

ข้อสอบถูกผิด (true-false) หรือ alternative-response items เป็นข้อคำถามที่กำหนดข้อความให้ผู้เรียนพิจารณาเลือกตอบ 2 ทางเลือก เช่น ถูกหรือผิด ใช่หรือไม่ใช่ จริงหรือไม่จริง เห็นด้วยหรือไม่เห็นด้วย เป็นต้น โดยใช้ความรู้ตามหลักวิชาเป็นเกณฑ์พิจารณาสามารถวัดได้ถึงระดับการนำไปใช้ มีรายละเอียด ดังนี้

หลักการสร้างข้อสอบแบบถูกผิด

1) แต่ละข้อความควรถามจุดสำคัญเพียงเรื่องเดียว หลีกเลี่ยงการใช้ส่วนขยายที่ก่อให้เกิดความสับสน

2) ข้อความสั้น กระชับ แต่มีความหมายชัดเจนไม่กำกวมใช้คำศัพท์และโครงสร้างประโยคง่ายๆ

3) แต่ละข้อความสามารถตัดสินได้ชัดเจนว่าถูกหรือผิด หลีกเลี่ยงการใช้คำที่แสดงคุณภาพ เช่น มาก น้อย บ่อย ๆ บางครั้ง ส่วนมาก ส่วนน้อย ไม่ค่อยจะ เท่านั้น เป็นต้น ควรเลือกคำที่แสดงปริมาณจะมีความหมายชัดเจน

4) ใช้ข้อความเชิงปฏิเสธเฉพาะที่จำเป็นเท่านั้น และควรขีดเส้นใต้หรือเน้นตัวหนาที่คำเชิงปฏิเสธให้ชัดเจน

5) ไม่ควรสร้างข้อความเชิงปฏิเสธหรือปฏิเสธซ้อนเพราะจะทำให้ผู้สอบเข้าใจผิด

6) ควรหลีกเลี่ยงการลอกข้อความจากหนังสือหรือตำราเรียนโดยตรงเพราะจะส่งเสริมการเรียนรู้แบบท่องจำ

7) ให้ข้อความแต่ละข้อเป็นอิสระแก่กัน

8) ข้อความแต่ละข้อควรมีความยาวใกล้เคียงกัน

9) ข้อสอบควรเรียงลำดับตามเนื้อหา

10) ควรให้มีจำนวนข้อถูกและข้อผิดใกล้เคียงกันและอยู่กระจายคละกัน

ข้อดีและข้อจำกัดของข้อสอบแบบถูกผิด

ข้อดี สร้างง่ายรวดเร็ว สะดวกต่อการตรวจให้คะแนน มีความเป็นปรนัยในการให้คะแนน สูงเหมาะกับการตรวจสอบความเข้าใจในแต่ละหัวข้อที่เรียน เหมาะกับเด็กเล็กที่ยังอ่านหนังสือไม่คล่อง

ข้อจำกัด สร้างให้มีคุณภาพได้ค่อนข้างยากผู้สอบมีโอกาสตอบถูกต้องได้ร้อยละ 50 การที่ผู้สอบตอบถูกอาจเกิดจากการเดา มีค่าความเที่ยงต่ำ จึงทำให้ผลการสอบไม่สะท้อนความสามารถของผู้เรียนอย่างแท้จริง ไม่เหมาะสำหรับสร้างเป็นแบบสอบวินิจฉัย เนื่องจากไม่สามารถวินิจฉัยได้ว่าตอบผิดเพราะเหตุใด

2) ข้อสอบจับคู่ (matching)

ข้อสอบจับคู่มีลักษณะเป็นชุดของคำ ข้อความ รูปภาพ หรือสัญลักษณ์ที่เป็นคำตอบคู่กับข้อความแต่ละข้อซึ่งแบ่งออกเป็น 2 คอลัมน์ โดยทั่วไปคอลัมน์ทางซ้ายมือจะเป็นข้อความ หรือที่เรียกว่า "ตัวยีน" และคอลัมน์ขวามือคือ คำตอบ หรือ "ตัวเลือก" การจับคู่นี้มีลักษณะคล้ายกับข้อสอบแบบเลือกตอบหลายตัวเลือก โดยส่วนใหญ่จะกำหนดให้มีตัวเลือกจำนวนมากกว่า

ข้อสอบจับคู่เป็นข้อสอบที่เหมาะสมกับการวัดความรู้ความเข้าใจเกี่ยวกับความสัมพันธ์ หรือความเชื่อมโยงของสิ่งต่าง ๆ สามารถนำไปใช้ในหลากหลายวิชา หลากหลายเนื้อหาได้

หลักการสร้างข้อสอบจับคู่

1) เขียนคำชี้แจงให้ชัดเจนว่าจะให้จับคู่ได้เพียงตัวเลือกเดียวหรืออาจจับคู่ได้หลายตัวเลือก แต่ในกรณีที่ต้องการให้ตัวเลือกสามารถตอบถูกได้มากกว่า 1 ครั้ง ควรสร้างข้อสอบแบบเลือกตอบหลายตัวเลือกแบบคำถามเป็นชุดแต่ตัวเลือกคงที่

2) ตัวเลือกจะต้องมีจำนวนมากว่าตัวยีน 2-4 ข้อ กล่าวคือ ควรมีตัวลวงประมาณ 30% เนื่องจากหากมีจำนวนเท่ากันจะทำให้โอกาสที่จะเดาคำตอบถูกในข้อหลัง ๆ มีสูงมาก โดยเฉพาะข้อสุดท้ายที่สามารถตอบได้ทันทีโดยไม่ต้องคิดหาคำตอบ

3) ตัวยีนควรมีจำนวน 5-15 ข้อ หากน้อยเกินไปโอกาสที่จะเดาถูกจะมีมากขึ้น ในขณะที่มีตัวยีนมากเกินไปผู้สอบจะเกิดความสับสน การจับคู่จะยากเกินไป

4) เนื้อหาวิชาที่นำมาออกข้อสอบจะต้องเป็นเรื่องหรือเนื้อหาเดียวกันคำหรือข้อความที่เป็นคู่กันไม่ควรจัดให้อยู่ตรงกัน และควรจัดตัวเลือกแบบผสมผสานไม่มีเรียงลำดับอย่างเป็นระบบ

5) ต้องระบุความสัมพันธ์ของข้อความทั้งสองชุดให้ชัดเจน โดยชี้แจงให้ชัดเจนว่าจับคู่โดยยึดความสัมพันธ์แบบใด

ข้อดีและข้อจำกัดของข้อสอบจับคู่

ข้อดี สร้างง่ายรวดเร็ว สะดวกต่อการตรวจให้คะแนน มีความเป็นปรนัยในการให้คะแนน สูงเหมาะกับการวัดเนื้อหาที่เป็นการวัดความเข้าใจในความสัมพันธ์ของสิ่งต่างๆ ข้อสอบสามารถออกแบบให้ใช้รูปภาพ แผนภูมิ กราฟได้ ลดการเดาได้มาก และใช้เวลาในการทดสอบน้อย

ข้อเสีย ไม่สามารถวัดความรู้ความคิดระดับสูง ผู้สอนที่ออกข้อสอบประเภทนี้มาก ผู้สอบมีแนวโน้มจะอ่านหนังสือสอบแบบท่องจำ ไม่เหมาะสมไปสร้างข้อสอบจำนวนมาก ๆ หรือวัดได้ครอบคลุมทุกเนื้อหา และโอกาสตอบถูกแต่ละข้อไม่เท่ากัน

3) ข้อสอบเลือกตอบ (multiple choice)

ข้อสอบเลือกตอบ (multiple choice) หรือแบบสอบหลายตัวเลือก เป็นแบบสอบที่ได้รับค่านิยมต่อเนื่องมาจนถึงปัจจุบัน เนื่องจากเป็นแบบสอบที่สามารถวัดระดับความรู้ได้สูงกว่าแบบสอบประเภทเลือกคำตอบประเภทอื่น ๆ สามารถสร้างให้วัดครอบคลุมเนื้อเรื่องตามโครงสร้างอย่างมีประสิทธิภาพ

ข้อสอบแบบเลือกตอบประกอบด้วย 2 ส่วน ได้แก่ (1) ข้อคำถาม (stem) หรือประโยคที่ไม่สมบูรณ์ และ (2) ตัวเลือก (choices/alternatives/options) หรือข้อความที่เติมเต็มให้ประโยคมีความถูกต้องสมบูรณ์ถูกต้อง ซึ่งนิยม 3-6 ตัวเลือก ตัวเลือกนี้จะประกอบด้วยคำตอบที่ถูกต้อง เรียกว่า ตัวคำตอบ (answer/key) 1 ตัว และส่วนที่เลือกเป็นตัวเลือกที่ผิด หรือที่เรียกว่า ตัวลวง (distracters)

รูปแบบคำถามของข้อสอบเลือกตอบ

รูปแบบคำถามของข้อสอบเลือกตอบสามารถแบ่งออกได้เป็น 3 ประเภท คือ 1) แบบคำถามเดี่ยว 2) แบบคำถามเป็นชุดแต่ตัวเลือกคงที่ และ 3) แบบสถานการณ์ แต่ละแบบมีรายละเอียดดังนี้ (โชติกา ภาชีผล และคณะ, 2558; พิชิต ฤทธิ์จรูญ, 2556; สมนึก ภัททิยธนี, 2560)

1) แบบคำถามเดี่ยว (single question) คำถามลักษณะนี้เป็นที่นิยมมาก เป็นคำถามที่มีความสมบูรณ์ในแต่ละข้อนั้น ๆ และเป็นอิสระจากข้อคำถามอื่น มี 14 แบบ ดังนี้

1.1) แบบคำตอบถูก เป็นแบบกำหนดให้มีคำตอบที่ถูกต้องเพียงข้อเดียว มี 3 ลักษณะ คือ คำตอบถูกต้อง คำตอบที่ดีที่สุด และคำตอบใกล้เคียง ซึ่งคำตอบใกล้เคียงนี้แบ่งได้อีก 3 ชนิด คือ ภาษาคำใกล้เคียง ภาษ้อัตราส่วน และภาษเป็นพหุหรือขอบเขต

1.2) แบบเติมคำ ดัดแปลงมาจากข้อสอบแบบเติมคำ แบ่งเป็น การเติมคำเพียงแห่งเดียว และการเติมคำสองแห่ง

1.3) แบบเปลี่ยนแทน เป็นข้อสอบที่ให้ผู้สอบหาคำหรือวลีใหม่มาเปลี่ยนแทนถ้อยคำที่ไม่สมบูรณ์ มี 2 ลักษณะ คือ แบบเปลี่ยนแปลง เป็นข้อสอบที่มุ่งวัดความสามารถในการแก้ไขสิ่งที่บกพร่องหรือผิดพลาดให้ถูกต้อง และแบบปรับปรุง คือมุ่งให้ผู้สอบปรับปรุงถ้อยคำหรือสำนวนใหม่ให้สละสลวยและสมบูรณ์มากยิ่งขึ้น

1.4) แบบคำตอบคู่ เป็นข้อสอบที่ให้ผู้สอบหาคำตอบที่ดีที่สุดควบคู่กัน

1.5) แบบคำตอบผสม หรือคำตอบซ้อนเป็นข้อสอบที่ข้อคำถามเขียนเป็นเงื่อนไข ซึ่งควรมีอย่างน้อย 3 เงื่อนไข แล้วเขียนตัวเลือกจากเงื่อนไขที่กำหนดไว้เพื่อให้ผู้สอบได้พิจารณาจากเงื่อนไขหลาย ๆ ตัว

1.6) แบบคำตอบไม่สมบูรณ์ เป็นข้อสอบที่ตัวเลือกที่กำหนดให้ยังไม่ใช่คำตอบที่สมบูรณ์ ผู้สอบจำต้องคิดหาคำตอบจากคำตอบที่กำหนดให้อีกทีหนึ่ง

1.7) แบบเรียงอันดับเป็นข้อสอบที่ต้องการให้ผู้สอบเรียงลำดับ เช่น ตามความสำคัญ ตามขั้นตอนก่อนหลัง ตามเหตุการณ์ หรือเวลา เป็นต้น

1.8) แบบจำแนกประเภท เป็นข้อสอบที่ต้องการให้ผู้สอบจำแนกประเภทชนิดที่เข้าพวก ชนิดที่ต่างจากพวก และชนิดเชื่อมโยง

1.9) แบบสัมพันธ์ เป็นข้อสอบที่ต้องการให้ผู้สอบหาความสัมพันธ์ของสองสิ่ง หรือสองเรื่องขึ้นไป ความสัมพันธ์เช่น เหตุและผล อุปมาอุปมัย เป็นต้น

1.10) แบบขาดเกิน เป็นข้อสอบที่ต้องการให้ผู้สอบวินิจฉัยความสมบูรณ์ของเรื่องราวว่า ยังมีสิ่งที่บกพร่อง หรือสิ่งที่เกินมาโดยไม่จำเป็น เรียกว่า ชนิดขาด ชนิดเกิน และชนิดเพียงพอ

1.11) แบบหาตัวร่วม-ตัวต่าง เป็นข้อสอบที่ต้องการให้ผู้สอบหาสาระสำคัญ หรือแก่นของสิ่งนั้นมีลักษณะร่วมกันหรือต่างกันอยู่

1.12) แบบอนุกรม เป็นข้อสอบที่ต้องการให้ผู้สอบหารูปแบบ หรือกฎเกณฑ์จากโจทย์เพื่อหาคำตอบ ได้แก่ ชนิดอนุกรม และอนุกรมสัมพันธ์

1.13) แบบสรุปเรื่องราว เป็นข้อสอบที่ต้องการให้ผู้สอบพิจารณาข้อมูลหรือสิ่งที่โจทย์กำหนดให้ แล้วสรุปอย่างมีเหตุผล

1.14) แบบรูปภาพ เป็นข้อสอบที่ต้องการให้ผู้สอบพิจารณารูปภาพ เครื่องหมาย หรือสัญลักษณ์เพื่อหาคำตอบ

2) แบบคำถามเป็นชุดแต่ตัวเลือกคงที่ (constant choice question) เป็นข้อสอบที่ใช้ข้อความหลายข้อ แต่ใช้ชุดตัวเลือกชุดเดียวกัน มีลักษณะคล้ายข้อสอบแบบจับคู่ แต่มีความสะดวกและมีประสิทธิภาพมากกว่า เนื่องจากสามารถใช้ตัวเลือกใดๆ เป็นคำตอบซ้ำได้ และจะมีข้อความที่ซ้ำก็ได้ โดยที่มีตัวเลือกเพียง 4-5 ตัวเลือก แต่อย่างไรก็ดี การกำหนดค่าชี้แจงเป็นเรื่องสำคัญที่จะต้องแจ้งให้ผู้สอบทราบว่าตัวเลือกใช้สำหรับข้อความใด และต้องการให้พิจารณาคำตอบในแง่มุมใด

3) แบบสถานการณ์ (situation) สถานการณ์อาจอยู่ในรูปแบบของข้อความ แผนภาพ แผนภูมิ บทความ เรื่องราว การตั้งคำถามจะไม่ถามให้ตรงเรื่องหรือถามนอกเรื่อง แต่ควรถามให้เกี่ยวข้องกับสัมพันธ์กับเรื่องราวนั้น

หลักการสร้างข้อสอบเลือกตอบ

ข้อคำถาม

- 1) ข้อสอบแต่ละข้อมุ่งวัดผลการเรียนรู้ที่สำคัญ
- 2) ควรเขียนข้อคำถามให้เป็นประโยคคำถามสมบูรณ์ อาจใส่เครื่องหมายปริศน์ (?)
- 3) เน้นเรื่องที่จะถามให้ชัดเจน ตรงจุดไม่คลุมเครือไม่ใช่คำพุ่มเพื่อยเขียนให้ถูกหลักไวยากรณ์ และหลีกเลี่ยงคำที่ขึ้นนำคำตอบ
- 4) ควรถามในเรื่องที่มีคุณค่าต่อการวัด หรือถามในสิ่งที่มีประโยชน์
- 5) ใช้ภาษาให้เหมาะสมกับระดับผู้สอบ
- 6) คำถามควรเร้าให้ผู้สอบได้ใช้ความคิด
- 7) หากเป็นข้อคำถามทางลบ เช่น **ไม่ ยกเว้น ไม่ใช่ ไม่ถูกต้อง** ให้เน้นคำหรือข้อความนั้น โดยการทำตัวหนา ตัวเอียง หรือขีดเส้นใต้หากเป็นภาษาอังกฤษสามารถใช้เป็นตัวพิมพ์ใหญ่ได้
- 8) หลีกเลี่ยงการใช้คำถามปฏิเสธซ้อน
- 9) ไม่ควรถามในสิ่งที่เด็กท่องจำจนคล่องปาก
- 10) คำถามแต่ละข้อควรเป็นอิสระขาดจากกัน

ตัวเลือก

- 1) คำถามข้อหนึ่ง ๆ ต้องมีคำตอบที่ถูกต้องที่สุดเพียงตัวเลือกเดียวเท่านั้น
- 2) เขียนให้ทั้งตัวถูกและตัวผิด ถูกผิดตามหลักวิชา
- 3) เขียนให้ตัวเลือกเป็นอิสระจากกัน

- 4) ตัวเลือกที่ถูกไม่ควรยาวเกินไป เขียนตัวเลือกให้กะทัดรัดไม่ยืดยาวหรือเพิ่มคำที่ไม่จำเป็น
- 5) ตัวเลือกต้องเป็นเอกพจน์
- 6) จัดตัวเลือกให้เป็นระบบ เช่น เรียงตาม พ.ศ. เรียงจากน้อยไปมาก เป็นต้น
- 7) หลีกเลี่ยงการเขียนตัวถูกให้ฟังเสียงหรือมีคำ/ข้อความที่ซ้ำกับตัวคำถาม
- 8) ตำแหน่งของตัวถูกควรกระจายในลักษณะสุ่ม
- 9) ตัวลวงต้องมีโอกาสเป็นไปได้
- 10) ไม่ควรมีตัวเลือกประเภท “ถูกหมดทุกข้อ” หรือ “ไม่มีข้อใดถูก”
- 11) อย่าแนะนำคำตอบ เช่น ออกคำถามวัดซ้ำ ตัวข้อความและตัวเลือกใช้ซ้ำกันหรือเกี่ยวข้องกัน ตัวเลือกใช้ศัพท์หรือภาษาแปลกกว่าข้ออื่น เป็นต้น

ข้อดีและข้อจำกัดของข้อสอบเลือกตอบ

ข้อดี สามารถวัดผลการเรียนรู้ได้หลากหลายวิชา ทำให้ได้สารสนเทศเพื่อการวินิจฉัยความรู้ ความเข้าใจของผู้สอบ ตรวจให้คะแนนได้ง่าย สะดวก รวดเร็ว มีความเป็นปรนัย สามารถวิเคราะห์คุณภาพ ข้อสอบได้ง่ายกว่าแบบสอบอื่น ๆ ขจัดปัญหาเรื่องการอ่านลายมือของผู้สอบไม่ออก สามารถวินิจฉัย ข้อบกพร่องหรือความไม่เข้าใจเนื้อหาได้อย่างเป็นระบบ

ข้อเสีย สิ้นเปลืองค่าใช้จ่ายสูง ใช้เวลาในการสร้างนานมาก โดยเฉพาะการสร้างให้ตัวลวงมีคุณภาพ การวัดความรู้ความสามารถระดับสูงสร้างได้ยาก

1.2 แบบสอบประเภทเสนอคำตอบ (supply type)

แบบสอบประเภทนี้เป็นแบบสอบที่ผู้สอบจะต้องอ่านคำถาม กำหนดแนวทางคำตอบ และเขียนคำตอบด้วยตนเอง โดยมีความยาวของคำตอบที่แตกต่างกันไป แบ่งออกเป็น ข้อสอบความเรียง (essay items) ข้อสอบแบบตอบสั้น (short answer items) หรือ ข้อสอบเติมคำ (completion items) แบบสอบประเภทเสนอคำตอบแต่ละประเภทมีคุณลักษณะ หลักการสร้าง และข้อดีและข้อจำกัดแตกต่างกัน ดังรายละเอียด ดังต่อไปนี้

1) อัตนัยหรือความเรียง (subjective or essay tests)

มีลักษณะเด่นที่ให้อิสระแก่ผู้ตอบ ผู้ตอบจะต้องหาหรือสร้างคำตอบเอง แทนที่จะมีคำตอบ ให้เลือกเหมือนกับข้อสอบแบบกำหนดคำตอบให้ ข้อสอบแบบนี้จะใช้ในการวัดผลการเรียนรู้ที่ไม่สามารถวัด โดยใช้ข้อสอบแบบกำหนดคำตอบได้ เช่น วัดความสามารถในการจัดการ ความสามารถในการบูรณาการ ความสามารถในการสังเคราะห์ ความรู้ความสามารถในการแก้ปัญหา ตลอดจนความสามารถในการ ประเมิน เป็นต้น เหมาะที่จะนำมาใช้เมื่อต้องการวัดความสามารถในการใช้เหตุผล การวางแผน การแสดง ความคิดเห็นสร้างสรรค์ หรือจินตนาการ กล่าวคือ เป็นการวัดผลการเรียนรู้ในระดับสูงและซับซ้อน ผู้ตอบ จะต้องรู้สึกในเรื่องที่จะตอบจึงจะเขียนตอบได้ดี ดังนั้น ข้อคำถามหรือคำสั่งจะต้องมีความชัดเจนว่าต้องการ ให้ตอบในเรื่องใดบ้าง

ข้อสอบแบบความเรียง แบ่งออกเป็น 2 ลักษณะ คือ คำถามแบบจำกัดคำตอบ (restricted response questions) เป็นคำถามที่จำกัดเนื้อหา คือ กำหนดขอบเขตที่จะให้ตอบ รวมถึงความยาวของ คำตอบและคำถามแบบขยายคำตอบ (extended response questions) เป็นคำถามที่ไม่จำกัดขอบเขตให้ ตอบผู้ตอบสามารถเลือกข้อความรู้และนำความรู้เหล่านั้นมาจัดระบบให้ดี แล้วนำมาเขียนเป็นคำตอบ จึง เป็นข้อสอบที่ใช้วัดความสามารถในการรวบรวมความรู้ สังเคราะห์ความรู้เหล่านั้นแล้วนำมาเรียบเรียงและ เขียนเป็นคำตอบตามความคิดและเหตุผลของตน

หลักการสร้างข้อสอบความเรียง

1. กำหนดให้ชัดเจนว่าต้องการวัดพฤติกรรมด้านใดของผู้สอบ
2. เขียนข้อความให้อยู่ในรูปประโยคคำถามให้ชัดเจนจำเพาะเจาะจงว่าต้องการให้ผู้สอบทำอย่างไรมีคะแนนข้อละกี่คะแนน
3. เขียนคำถามโดยใช้สถานการณ์ใหม่ ๆ ไม่ควรถามตามหนังสือหรือถามในสิ่งที่เรียนมาแล้ว
4. ต้องถามเฉพาะสิ่งที่เป็นประเด็นสำคัญของเรื่อง
5. กำหนดเวลาในการสอบให้เหมาะสมเพื่อผู้ตอบจะได้วางแผนการตอบได้ถูกต้อง โดยเอาจำนวนข้อไปหารจำนวนเวลาทั้งหมดก็จะทราบว่าแต่ละข้อควรใช้เวลาเท่าไร
6. ถ้าไม่จำเป็นควรหลีกเลี่ยงการให้ผู้ตอบเลือกตอบเป็นบางข้อ ควรให้ทำทุกข้อ
7. พยายามใช้คำถามหลาย ๆ แบบโดยเน้นการอธิบายควรเป็นคำถามประเภททำไมอย่างไรหรือให้อธิบาย บรรยาย เปรียบเทียบ หาความสัมพันธ์ หาความขัดแย้ง ตีความ วิเคราะห์เหตุผล วิจาร์ณและประเมินผลเป็นต้น
8. ควรเขียนคำตอบหรือแนวคำตอบที่ต้องการไว้ด้วยหรืออาจจะเขียนในลักษณะคำหรือข้อความสำคัญ (Key Words) ของคำตอบข้อนั้นๆ เอาไว้ด้วย
9. ควรกำหนดความยาวและความซับซ้อนของข้อสอบให้พอเหมาะกับความสามารถของผู้สอบ
10. ถ้าข้อสอบมีหลายข้อควรจัดเรียงลำดับจากง่ายไปหายากเพื่อยั่วให้ผู้อยากตอบมากยิ่งขึ้น
11. การเก็บข้อสอบและการตรวจให้คะแนนควรดำเนินการให้เสร็จเป็นครั้ง ๆ ไป โดยเฉพาะอย่างยิ่งการตรวจให้คะแนน ควรตรวจข้อนั้นๆ ให้เสร็จในเวลาเดียวกัน ไม่ทิ้งค้างการตรวจเป็นระยะ ซึ่งอาจทำให้ขาดความคงที่ของการให้คะแนน

ข้อดีและข้อจำกัดของข้อสอบความเรียง

ข้อดี สามารถใช้วัดได้ในทุกสาขาวิชา สามารถสร้างข้อสอบได้ในเวลาสั้น ๆ ข้อสอบช่วยส่งเสริมการคิดวิเคราะห์และสังเคราะห์ข้อมูล โอกาสในการเดามีน้อย ทำให้ทราบว่าผู้เรียนมีมีโนทัศน์ที่คลาดเคลื่อนในเนื้อหาที่เรียนเรื่องใดบ้าง

ข้อจำกัด สร้างข้อสอบเพื่อวัดได้จำนวนไม่มากนัก เนื่องจากต้องใช้เวลาในการตอบแต่ละข้อค่อนข้างมาก จึงวัดไม่ได้ครอบคลุมหลักสูตรหรือเนื้อหาสำคัญ การตรวจใช้เวลาานจึงไม่เหมาะกับการสอบกับนักเรียนจำนวนมากๆ และมีความคลาดเคลื่อนมาก ควบคุมให้เกิดความยุติธรรมได้ยาก ปลายมือของผู้สอบและประสิทธิภาพในการเขียนบรรยายอาจมีผลต่อคะแนนได้

2) ข้อสอบแบบตอบสั้น (short answer items) หรือ ข้อสอบเติมคำ (completion items)

เป็นข้อสอบที่เหมาะสมกับการใช้ตรวจสอบความรู้ความจำ มากกว่าความเข้าใจ เช่น การถามเกี่ยวกับคำศัพท์ จำนวน หรือสัญลักษณ์ เพื่อตอบคำถาม หรือทำให้เป็นประโยคที่สมบูรณ์ สามารถใช้ตรวจสอบการสะกดคำได้ถูกต้อง เหมาะสมกับเนื้อหาที่มีคำศัพท์เฉพาะ เป็นความรู้ที่เป็นจริงไม่ใช่ข้อคิดเห็น สามารถนำมาใช้

หลักการสร้างข้อสอบแบบตอบสั้น

1. ควรมีการวางแผนเนื้อหาที่จะวัดเรื่องคำศัพท์เฉพาะสาขา ความรู้ความจริง (fact) หรือความรู้พื้นฐานอื่น ๆ

2. พิจารณาถึงคำตอบที่เป็นไปได้ ซึ่งควรมีคำตอบเดียวเท่านั้นที่ถูกต้อง
3. พยายามใช้ข้อความที่มีสารสนเทศเพียงพอที่จะสามารถตอบข้อความได้ถูกต้อง
4. เว้นที่ว่างให้พอเหมาะกับคำตอบ และไม่สั้นหรือยาวจนเกินไป
5. ทำรายการคำตอบที่เป็นไปได้
6. ข้อคำถามไม่ควรใช้ข้อความจากตำรา ผู้สร้างข้อสอบควรเขียนขึ้นเอง
7. ข้อคำถามควรมีความชัดเจนมากที่สุดเท่าที่จะเป็นไปได้
8. หากข้อคำถามเป็นประโยคไม่สมบูรณ์ ควรเว้นช่องว่างสำหรับคำตอบไว้ท้ายข้อความ
9. เมื่อข้อคำถามมีค่าที่เป็นปฏิเสธ ควรเน้นข้อความนั้นให้ชัดเจน โดยขีดเส้นใต้ ทำเป็นตัว

ใหญ่ ตัวเอน หรือตัวทึบ

ข้อดีและข้อจำกัดของข้อสอบแบบตอบสั้น

ข้อดี สร้างได้ง่าย ลดปัญหาการเดา ช่วยจำแนกผู้สอบที่รอบรู้และไม่รอบรู้ออกจากกันได้ และสามารถนำมาสร้างเป็นข้อสอบอัตโนมัติยาวเป็นแบบย่อหน้าได้

ข้อจำกัด ส่วนใหญ่เป็นการวัดความรู้ความจำ การให้คะแนนอาจขึ้นอยู่กับผู้ตรวจ และอาจมีปัญหาการสะกดคำผิดได้

2. แบบวัดเชิงสถานการณ์

แบบวัดเชิงสถานการณ์ หมายถึง แบบทดสอบที่กำหนดให้มีสถานการณ์ที่คล้ายคลึง หรือเลียนแบบเหตุการณ์ต่าง ๆ ในชีวิตจริง เพื่อให้ผู้สอบได้สัมผัสกับสภาพการณ์เหล่านั้น แล้วแสดงพฤติกรรมตอบสนองที่สังเกตได้ หรือวัดได้ออกมา

สมนึก ภักดิ์ธยานี (2560) ได้กล่าวถึงแบบวัดเชิงสถานการณ์ว่าเป็นข้อสอบที่ใช้วิธีการกำหนดข้อความ ภาพ ตาราง ให้นักเรียนอ่านหรือพิจารณาดูก่อน แล้วตั้งคำถามเกี่ยวกับข้อความหรือภาพหรือตารางที่กำหนดให้

สมบูรณ์ ชิตพงศ์ (ม.ป.ป.) ได้ให้ความหมายของแบบวัดเชิงสถานการณ์ว่า เป็นการจำลองหรือสร้างเหตุการณ์เรื่องราวต่าง ๆ ขึ้น แล้วให้บุคคลแสดงความรู้สึกว่าตนเองจะกระทำ หรือมีความเห็นอย่างไร ต่อสถานการณ์ที่กำหนดขึ้น โดยปกติแล้วการตอบสนองต่อสถานการณ์นั้น อาจให้ตอบสนองว่าตัวเขาเอง จะทำอย่างไร หรือการให้เขาแสดงความคิดเห็นว่าตัวบุคคลในสถานการณ์นั้นๆ จะทำอย่างไร

จากการศึกษาความหมาย แนวคิดเกี่ยวกับแบบวัดเชิงสถานการณ์ สามารถสรุปได้ว่า แบบวัดเชิงสถานการณ์ หมายถึง แบบทดสอบที่มีการกำหนดสถานการณ์ขึ้นให้คล้ายกับเหตุการณ์ต่าง ๆ ในชีวิตจริง อาจเป็นข้อความ ภาพ หรือตารางตาราง เพื่อให้ผู้สอบได้สัมผัสสถานการณ์แล้วแสดงความรู้สึก ความคิดเห็น หรือพฤติกรรมที่วัดได้ออกมา

2.1 ลักษณะของแบบวัดเชิงสถานการณ์

1. เป็นแบบสถานการณ์มาให้แล้วถามความคิดเห็นของผู้ตอบเกี่ยวกับการกระทำของตัวละครในสถานการณ์ ว่าเห็นด้วยหรือไม่ ถ้าหากเป็นผู้ตอบจะทำเหมือนตัวละครในสถานการณ์นั้นหรือไม่

2) กำหนดสถานการณ์พร้อมทั้งกำหนดทางเลือกมาให้ 3 – 4 แนวทาง แล้วให้ผู้ตอบเลือกตอบ

3) ถามแนวทางประพจน์หรือปฏิบัติกิจกรรมเรื่องราวต่าง ๆ ตามที่กำหนดให้ เป็นการถามพฤติกรรมตรง ๆ ว่า ผู้ตอบเคยปฏิบัติมากน้อยเพียงใดในเหตุการณ์หรือสถานการณ์ที่กำหนดให้ สถานการณ์ที่กำหนดขึ้นควรเกิดขึ้นในชีวิตจริงและเหมาะสมกับระดับของผู้เรียน

2.2 หลักและวิธีการสร้างแบบวัดเชิงสถานการณ์

สมนึก ภัทธิยธานี (2560) กล่าวถึงหลักสำคัญในการสร้างแบบวัดเชิงสถานการณ์ดังนี้

1) สถานการณ์ที่กำหนดอาจใช้ข้อความ คำพูด คำสนทนา หรือรูปภาพแสดงการทดลองทางวิทยาศาสตร์ หรือเป็นตารางตัวเลข สถิติ หรือกราฟเกี่ยวกับสิ่งหนึ่งสิ่งใดก็ได้ เพียงแต่เป็นสถานการณ์ที่มีความหมาย มีแง่มุมให้คิดพิจารณาและควรเป็นสถานการณ์ที่รัดกุม ไม่ยืดเยื้อเกินความจำเป็น เพื่อนำไปใช้เป็นหลักในการตอบคำถามต่อไป

2) การถามคำถามแง่มุมที่ต้องคิดและพิจารณา ไม่ควรถามตรงตามสถานการณ์ ที่กำหนดหรือถามนอกสถานการณ์จนเป็นเรื่องทั่วไป ซึ่งสามารถตอบได้โดยไม่ต้องใช้สถานการณ์ที่ให้มา

3) แบบวัดเชิงสถานการณ์นิยมถามในแง่การแปลความหมายตามวัตถุประสงค์ของสถานการณ์ ข้อคิดที่ได้ให้หาความสัมพันธ์ของสิ่งต่าง ๆ รวมทั้งการตั้งชื่อข้อความหรือภาพที่กำหนดให้ นอกจากนี้การเขียนตัวเลือกควรหลีกเลี่ยงการใช้ตัวเลือกที่มีคำตรงกับคำในข้อความแต่ถ้ามีความจำเป็นหรือหลีกเลี่ยงไม่ได้ก็ควรให้มีตัวเลือกหลาย ๆ ตัวที่ใช้คำหรือข้อความเพื่อให้ดูคล้าย ๆ กัน

พิชิต ฤทธิ์จรูญ (2556) กล่าวว่า หลักและวิธีการสร้างแบบวัดเชิงสถานการณ์มีแนวปฏิบัติ ดังนี้

1) กำหนดเนื้อหาและพฤติกรรมหรือคุณลักษณะที่ต้องการจะวัดให้ชัดเจน

2) เลือกข้อความหรือสถานการณ์ที่มีความยากพอเหมาะกับระดับชั้นของผู้เรียนและเนื้อเรื่องหรือสถานการณ์ที่ใช้ถามจะต้องไม่ลำเอียงต่อเด็กกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ

3) พยายามเขียนคำถามเพื่อถามใจความเนื้อหาหรือสถานการณ์นั้นตามพฤติกรรม หรือคุณลักษณะที่ต้องการจะวัด ซึ่งการเขียนสถานการณ์และการเขียนคำถามมีข้อควรคำนึงดังนี้

การเขียนสถานการณ์ มีหลักดังนี้

1) สถานการณ์ที่สร้างขึ้น ควรเลือกสถานการณ์ที่มีความเป็นไปได้ ที่จะเกิดขึ้นจริง ๆ กับบุคคลหรือกลุ่มตัวอย่างในขณะนั้น

2) ปัญหาในสถานการณ์ที่สร้างขึ้นหรือกำหนดขึ้น ควรมีความเข้มหรือรุนแรงในระดับกลาง ๆ ไม่สร้างความเครียดให้กับผู้ตอบจนเกินไป เพราะหากสร้างปัญหาที่มีความเข้มเกินไปอาจทำให้ผู้ตอบไขว้เขวได้ เช่น เขียนสถานการณ์ว่าแม่ป่วยหนัก และต้องการผ่าตัดในอีก 2 วัน หากหาเงินไม่ได้จะต้องตาย ตนเองไม่มีเงิน หากมีคนมาเสนอให้ไปขายบริการทางเพศ 1 คืน จะได้เงินจำนวนมากพอรักษา จัดว่าเป็นสถานการณ์ที่เข้มรุนแรงเกินไป อาจทำให้ผู้ตอบไขว้เขวได้ ความจริงไม่อยากขายบริการทางเพศ (เป็นคนประกอบสัมมาชีพ) แต่ต้องการตอบแทนบุญคุณพ่อแม่ (เป็นคนกตัญญู)

3) สาระสำคัญที่กำหนดให้ในสถานการณ์ จะต้องเพียงพอที่จะให้ผู้ตอบตัดสินใจเลือกทางปฏิบัติในแนวทางที่เหมาะสม

การเขียนคำถาม มีหลักดังนี้

1) ไม่ควรถามตรง ๆ แต่ควรถามให้เกี่ยวพันเรื่องราว – สถานการณ์ที่กำหนดไว้และไม่ควรถามนอกเรื่องที่ไม่ได้ใช้ข้อความในสถานการณ์นั้นมาช่วยตอบ หรือไม่ควรถามในกรณีถ้าไม่มีสถานการณ์นั้นแล้วก็สามารถตอบคำถามนั้นได้

2) ในการเลือกสถานการณ์เพื่อมาตั้งคำถามควรเลือกเฉพาะเนื้อหาหรือความรู้ที่เป็นตัวแทนที่มีความสำคัญต่อวิชานั้นมาถาม ไม่ควรนำเรื่องปลีกย่อยหรือรายละเอียดปลีกย่อยของรายวิชามาตั้งเป็นสถานการณ์ และไม่ควรถามด้วยการหลอกล่อให้ผู้ตอบตกหลุมด้วยเรื่องที่ไร้สาระ

3) คำถามที่ใช้อาจมี 2 ลักษณะ คือ

(1) ถามให้ประเมินสถานการณ์ดังกล่าวเพื่อตัดสินว่าควร-ไม่ควร ดี-ไม่ดี ทำ-ไม่ทำ ถูกต้อง-ไม่ถูกต้อง ใช้ได้-ใช้ไม่ได้ และรวมถึงกรณีที่ไม่อาจตัดสินใจได้ด้วย

(2) ถามให้ระบุแนวทางที่ตนเองจะปฏิบัติ ถ้าหากตนเองเป็นบุคคลในสถานการณ์นั้น หรือผู้เกี่ยวข้องในสถานการณ์นั้นจะปฏิบัติอย่างไร

4) เมื่อเขียนสถานการณ์และข้อคำถามเสร็จแล้ว ให้ทบทวนว่าสถานการณ์เหมาะสมเป็นปัจจุบันหรือไม่ สารระที่กำหนดไว้เพียงพอที่จะตัดสินใจได้หรือไม่

5) นำแบบวัดไปทดลองใช้และปรับปรุงแก้ไข

Ebel (1972 อ้างถึงใน ขวัญยืน มูลศรี, 2548) ได้กล่าวถึงการสร้างแบบวัดเชิงสถานการณ์ไว้ดังนี้

1) หลักการสร้างแบบวัดเชิงสถานการณ์มีแนวปฏิบัติ ดังนี้

(1) กำหนดเนื้อหา หรือพฤติกรรมหรือคุณลักษณะที่ต้องการจะวัด

(2) เลือกหัวข้อหรือสถานการณ์ที่จะใช้ถาม

(3) เขียนคำถามเพื่อถามตามใจความในเนื้อหาหรือสถานการณ์นั้น

ข้อดีของแบบวัดเชิงสถานการณ์ (พิชิต ฤทธิ์จรูญ, 2556)

1) แบบวัดเชิงสถานการณ์เป็นแบบวัดที่แสดงถึงฝีมือ หรือความสามารถของผู้เขียนข้อสอบว่าสามารถนำความรู้ที่เรียนมาผนวกกับเงื่อนไขในสถานการณ์ที่กำหนดได้ดีเพียงใด

2) สามารถวัดความรู้ขั้นสูงทั้งด้านสมรรถภาพทางสมอง และด้านจิตพิสัย

3) ระวังผู้ตอบให้ติดตามและอ่านเรื่องราวและได้คิดมากกว่าข้อสอบประเภทอื่น ๆ

4) สร้างความยุติธรรมให้แก่ผู้เข้าสอบทุกคน เพราะได้อ่านสถานการณ์เดียวกันทั้งหมด ไม่มีใครได้เปรียบเสียเปรียบเพราะใช้ตำราต่างกันหรือการสอนแตกต่างกัน เป็นต้น

ข้อจำกัดของแบบวัดเชิงสถานการณ์ (พิชิต ฤทธิ์จรูญ, 2556)

1) การเขียนคำชี้แจงของแบบวัดเชิงสถานการณ์ ต้องพึงระวังเป็นพิเศษ ต้องชี้แจงให้ผู้ตอบใช้สถานการณ์ที่กำหนดเป็นหลัก ถึงจะผิดแปลกจากความเป็นจริงก็ต้องตอบตามนั้น

2) สร้างค่อนข้างยุ่งยาก ผู้เขียนข้อสอบจะต้องเลือกสถานการณ์ที่เป็นปัจจุบัน ไม่เข้มจนเกินไป และจะต้องล้วงลึกเฉพาะในสถานการณ์ที่กำหนดให้เท่านั้น

3) กำหนดเกณฑ์ในการให้คะแนนค่อนข้างทำได้ยาก

3. มาตรประมาณค่า

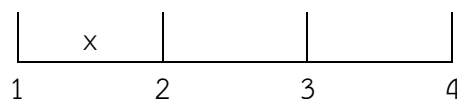
มาตรประมาณค่าคือแบบวัดที่ได้รับความนิยมมากรูปแบบหนึ่งโดยมีข้อความหรือประเด็นปัญหาให้ผู้ประเมินหรือผู้อ่านตัดสินใจเลือกตามความรู้สึกหรือความคิดเห็นของผู้ตอบ (ทิพย์ฉิมพร เกษโกมล, 2554 และ อนุวัติ คุณแก้ว, 2558)

3.1 รูปแบบและกระบวนการสร้างมาตรประมาณค่า

วรรณดี แสงประทีปทอง (2545) กล่าวว่า มาตรประมาณค่าเป็นมาตราที่ใช้ในการพิจารณาให้ค่าสิ่งเร้าโดยการพิจารณาตัดสินใจบนช่วงของความต่อเนื่องจากมากที่สุดไปถึ้น้อยที่สุด นิยมใช้ในการประเมินสิ่งเร้าเกี่ยวกับความรู้สึก พฤติกรรม กระบวนการและผลผลิต เช่นเดียวกับ Reynolds et al. (2009) ที่กล่าวว่า มาตรประมาณค่าใช้ในการตัดสินคุณภาพเพื่อบ่งชี้ความสามารถหรือการตัดสินใจที่รูปแบบของมาตรประมาณค่ามีส่วนของคำตอบเป็นตัวเลือกที่แสดงถึงความเข้มและ/หรือทิศทาง (สุวิมล ติรกานนท์, 2551) โดยมีคำตอบให้เลือก 3-8 ตัวเลือก ถ้าเป็น 3 ระดับคำตอบที่ให้เลือกได้แก่ เห็นด้วย ไม่

แน่ใจ ไม่เห็นด้วย ถ้าเป็น 5 ระดับ คำตอบที่ให้เลือกได้แก่ เห็นด้วยอย่างยิ่ง เห็นด้วยมาก เห็นด้วยปานกลาง เห็นด้วยน้อย และไม่เห็นด้วยอย่างยิ่ง เป็นต้น ทั้งนี้กระบวนการสร้างมาตรฐานค่ามีขั้นตอนดังนี้ (ทิพย์ คุ้มพร เกษโกมล, 2554 และ อนุวัติ คุณแก้ว, 2558)

1. กำหนดลักษณะของสิ่งที่ต้องการประเมิน
2. การกำหนดนิยามเชิงปฏิบัติการของตัวแปร เพื่อกำหนดสิ่งที่ต้องการวัดโดยได้มาจากแนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้อง
3. การสร้างตารางโครงสร้างของแบบวัด (test blueprint) เพื่อกำหนดองค์ประกอบหรือมิติของการวัด ทิศทางของข้อความ จำนวนข้อ
4. การเขียนข้อความหรือสร้างข้อคำถามจำนวนมากโดยเลือกข้อความที่สะท้อนจุดมุ่งหมายของการวัดและมีความเกี่ยวข้องกัน
5. การคัดเลือกข้อความ โดยเกณฑ์การคัดเลือกข้อความมีดังนี้
 - 5.1 ข้อความนั้นเกี่ยวข้องกับคุณลักษณะที่ต้องการวัด ซึ่งคาดว่าผู้ที่มีคุณลักษณะต่างกันจะตอบข้อความนั้นในทิศทางและขนาดที่ต่างกัน
 - 5.2 ข้อความกระชับ ชัดเจน
 - 5.3 ภาษายากง่ายเหมาะสมกับความเข้าใจของผู้ตอบ ไม่กำกวม ไม่มีหลายนัย
 - 5.4 ข้อความไม่เป็นปฏิเสธซ้อนที่เข้าใจยาก
 - 5.5 ข้อความนั้นมีใจความเดียว
 - 5.6 ข้อความควรมีทิศทางบวกและลบ และวางสลับแบบสุ่ม
 - 5.7 จำนวนข้อครอบคลุมทุกมิติที่วัด รวมเป็น 2 เท่าของจำนวนที่ใช้ในฉบับจริง
6. การกำหนดระดับการประมาณค่าแต่ละข้อความ โดยทั่วไปใช้ 5 ระดับ 7 ระดับ จากนั้นแบ่งเป็นสองช่วงแบบสมมาตรโดยมีทางบวกและลบฝ่ายละเท่ากันและมีจุดศูนย์อยู่กลาง ทั้งนี้การเขียนข้อความและการกำหนดรูปแบบของมาตรฐานค่ามีดังนี้ (Hopkins & Antes, 1990 อ้างถึงใน อนุวัติ คุณแก้ว, 2558)
 - 6.1 หลีกเลี่ยงการกำหนดคะแนนที่มีความละเอียดจนเกินไปหรือความห่างของคะแนนน้อยมาก ได้แก่ การกำหนดเป็นทศนิยม เป็นต้น
 - 6.2 หลีกเลี่ยงการกำหนดคะแนนที่ห่างกันมาก เช่น 1, 5, 10, 15, 20 เป็นต้น เพราะจะทำให้ยากในการจำแนก
 - 6.3 ควรใช้ระดับความเข้มหรือคะแนนที่เป็นจำนวนคู่ เช่น 4, 6, 8 ระดับ หรือ เห็นด้วยอย่างยิ่ง เห็นด้วย ไม่เห็นด้วย ไม่เห็นด้วยอย่างยิ่ง เป็นต้น
 - 6.4 ถ้ามาตราส่วนประมาณค่าเป็นแบบเส้นหรือกราฟ ควรให้ผู้ตอบทำเครื่องหมายที่ตรงกับเลขนั้น ๆ ถ้าให้ทำเครื่องหมายระหว่างตัวเลขจะทำให้การแปลผลกำกวมเพราะไม่รู้ว่าจะเลือก 1 หรือ 2 เช่น



6.5 ถ้าเป็นตัวเลขควรมีคำอธิบายว่า ตัวเลขแต่ละตัวแทนความหมายอะไร เช่น

- 4 แทน เห็นด้วยอย่างยิ่ง
 3 แทน เห็นด้วย
 2 แทน ไม่เห็นด้วย
 1 แทน ไม่เห็นด้วยอย่างยิ่ง

6.6 การกำหนดระดับความเข้มของความรู้สึกอาจใช้เป็นภาษา เป็นประโยค วลี คำก็ได้แต่ต้องเป็นคำที่มองเห็นถึงความแตกต่างกัน โดยมาตรฐานค่ามีรูปแบบหลายรูปแบบดังนี้

แบบที่ 1 แบบตัวเลข (numerical rating scales) เป็นการใช้ตัวเลขแทนระดับความคิดเห็น ดังนี้

ข้อความ	5	4	3	2	1

- 5 แทน เห็นด้วยอย่างยิ่ง
 4 แทน เห็นด้วย
 3 แทน ไม่แน่ใจ
 2 แทน ไม่เห็นด้วย
 1 แทน ไม่เห็นด้วยอย่างยิ่ง

แบบที่ 2 แบบบรรยาย (descriptive rating scales) คล้ายกับตัวเลข ต่างกันตรงที่ใช้เฉพาะข้อความบอกระดับความรู้สึก เช่น

ตัวอย่าง นักเรียนดูรายการข่าวโทรทัศน์มากน้อยเพียงใด

- ก. เป็นประจำ
 ข. ค่อนข้างบ่อย
 ค. พอประมาณ
 ง. นาน ๆ ครั้ง
 จ. ไม่เคยเลย

แบบที่ 3 แบบกราฟ (graphic rating scales) เป็นการใช้กราฟเส้นตรงแบ่งเขตระดับความรู้สึก ใช้ภาษา ตัวเลข หรือทั้งภาษาและตัวเลข ประกอบที่กราฟ เช่น

ตัวอย่าง การออกกำลังกายของนักเรียน

เป็นประจำ	5	4	3	2	1	ไม่เคยเลย
-----------	---	---	---	---	---	-----------

แบบที่ 4 แบบใช้สัญลักษณ์ (Symbolic rating scales) เป็นแบบที่ใช้สัญลักษณ์หรือรูปภาพเป็นคำตอบ เหมาะสำหรับเด็กเล็ก เช่น

ตัวอย่าง การเรียนคณิตศาสตร์



แบบที่ 5 แบบจัดลำดับที่ (Ranking) เป็นแบบให้ผู้ตอบจัดลำดับที่ หรือ การเรียงตามความสนใจ เช่น

ตัวอย่าง จัดลำดับวิชาที่ชอบมากที่สุด ไปหาน้อยที่สุด โดยให้ผู้ตอบใส่ตัวเลขตั้งแต่ 1 ถึง 5 ซึ่ง 1 หมายถึง น้อยที่สุด 5 หมายถึง มากที่สุด

_____ ภาษาอังกฤษ
 _____ กีฬา
 _____ ดนตรี
 _____ ศิลปะ
 _____ วิทยาศาสตร์

ข้อดีและข้อจำกัดของมาตรประมาณค่า

อนูวัตติ คุณแก้ว (2558) กล่าวถึงข้อดีและข้อจำกัดของมาตรประมาณค่า ดังนี้

ข้อดี

1. ประเมินได้ละเอียดกว่าแบบตรวจสอบรายการ
2. มีรูปแบบที่หลากหลายทำให้ผู้ตอบไม่เบื่อหน่าย
3. นำผลไปวิเคราะห์และแปลผลได้ง่ายเพราะมีตัวเลือกให้ตอบ

ข้อจำกัด

1. มีความคลาดเคลื่อนในการประเมิน (Rating Errors) ความคลาดเคลื่อนนี้เป็นความคลาดเคลื่อนที่เกิดจากผู้ประเมิน นำแบบประเมินซึ่งเป็นมาตรประมาณค่าไปประเมิน เช่น นำไปประเมินในขณะสังเกตพฤติกรรมของนักเรียน เป็นต้น โดยความคลาดเคลื่อนเกิดขึ้นได้ 4 กรณี (McIntire & Miller, 2000; McMillan, 2001 อ้างถึงใน อนูวัตติ คุณแก้ว, 2558) คือ

1.1 Leniency errors หรือ Generosity error เป็นความคลาดเคลื่อนที่เกิดจากผู้ประเมินให้คะแนนในการประเมินที่สูงกว่าความเป็นจริง

1.2 Severity errors เป็นความคลาดเคลื่อนที่เกิดจากผู้ประเมิน ให้คะแนนในการประเมินที่น้อยกว่าความเป็นจริง

1.3 Central-tendency error เป็นความคลาดเคลื่อนที่เกิดจากผู้ประเมินให้คะแนนตรงกลางเป็นส่วนใหญ่ ไม่ประเมินในส่วนที่มากที่สุดหรือน้อยที่สุด

1.4 Halo effect เป็นความคลาดเคลื่อนที่เกิดจากผู้ประเมินใช้การตัดสินใจจากองค์ประกอบหรือมิติเดียว ไม่พิจารณาสิ่งอื่นประกอบ เช่น นักเรียนคนหนึ่งได้รับการประเมินว่างานที่ส่งนั้นมีคุณภาพต่ำ ถึงแม้ว่าจำนวนงานที่ให้ทำนั้นนักเรียนคนนี้จะทำครบตามที่ส่งก็ตาม จะตัดสินใจว่าไม่ผ่านเกณฑ์การประเมินนั้นเลย หรือ ครูมีความชื่นชอบนักเรียนเป็นการส่วนตัวก็จะให้คะแนนนักเรียนคนนั้นสูงโดยไม่พิจารณาผลงานที่ทำเลย

2. ผู้ตอบไม่มีโอกาสแสดงความคิดเห็นเพราะต้องตอบตามตัวเลือกที่กำหนด

3. ถ้ามีระดับความเข้มของความรู้สึกแบ่งเป็นระดับคือ คือ แบ่งเป็น 3,5,7 ระดับผู้ตอบส่วนใหญ่ตอบระดับที่เป็นกลาง

ตามที่กล่าวมาข้างต้นจะเห็นได้ว่า มาตรประมาณค่าสามารถมีได้หลายระดับ ตั้งแต่ 3 ระดับ, 5 ระดับตามแนวคิดของลิเคิร์ท, 7 ระดับตามแนวคิดของออสกูด, และ 11 ระดับตามแนวคิดของเทอร์สตัน ทั้งนี้มาตรประมาณค่าของออสกูดและเทอร์สตันนั้นมุ่งเน้นการวัดเป้าหมายการเรียนรู้ด้านเจตคติ ในขณะที่มาตรวัดตามแนวคิดของลิเคิร์ทนิยมนำมาประยุกต์ใช้ทั้งด้านเจตคติและระดับพฤติกรรม ประกอบกับมีงานวิจัยมากมายที่ศึกษาความสัมพันธ์เกี่ยวกับระดับคุณภาพของเครื่องมือและรูปแบบของมาตรประมาณค่าที่พบว่ามาตรวัดตามแนวคิดลิเคิร์ท 5 ระดับนั้นให้คุณภาพของเครื่องมือสูงกว่าออสกูดและเทอร์สตัน

3.2 หลักการสร้างมาตรฐานค่าตามแนวคิดของลิเคิร์ต

สุวิมล ติรภานันท์ (2551) กล่าวว่า Likert คิดวิธีการนี้ในปี ค.ศ.1932 เป็นการสร้างข้อคำถามขึ้นจำนวนหนึ่งซึ่งมีลักษณะเป็นตัวบ่งชี้รวมเพื่อวัดเรื่องใดเรื่องหนึ่งต่อสิ่งใดสิ่งหนึ่ง คะแนนที่ได้ในแต่ละข้ออยู่ในมาตรฐานบัญญัติ จากนั้นนำคะแนนที่ได้จากข้อความทั้งหมดคะแนนรวมที่ได้จะอยู่ในมาตรฐานอัตราภาคหลักการในการสร้างมาตรวัดของลิเคอร์ทมีดังนี้

1. ข้อความที่เป็นคำถามมีทั้งข้อความที่เป็นทางบวกและทางลบเพื่อมิให้ผู้ตอบเกิดความเคยชินในการตอบ แต่เดิมมีข้อกำหนดว่าควรมีจำนวนใกล้เคียงกัน แต่งานวิจัยในปัจจุบันพบว่าสัดส่วนข้อความทางบวกและลบไม่ว่าจะจะเป็น 50:50 หรือ 75:25 หรือ 25:75 ให้ผลไม่แตกต่างกัน และในบางกรณีพบว่าการใช้ข้อความทางลบทั้งฉบับกลับทำให้ค่าความเที่ยงสูงกว่าการใช้ข้อความทางบวกและทางลบในฉบับเดียวกัน (พิบูล กิจจศิริ, 2543 อ้างถึงใน สุวิมล ติรภานันท์, 2551) ในการสร้างข้อคำถามไม่ควรสร้างข้อความทางบวกและทางลบในเนื้อหาเดียวกันเพราะจะเป็นการถามซ้ำและทำให้ข้อความนี้มีคะแนนเป็น 2 เท่าของที่ควรจะเป็น

2. การให้คะแนนต้องเป็นทางเดียวกับลักษณะของข้อความ

	ข้อความทางบวก	ข้อความทางลบ
เห็นด้วยอย่างยิ่ง	5	1
เห็นด้วย	4	2
ไม่แน่ใจ	3	3
ไม่เห็นด้วย	2	4
ไม่เห็นด้วยอย่างยิ่ง	1	5

การให้คะแนน 5 4 3 2 1 ของแต่ละข้อความเกิดมาจากการสร้างมาตรวัดทางจิตวิทยา ซึ่งพบว่าคะแนนที่ได้จากการสร้างทุกครั้งจะได้คะแนนใกล้เคียงกับ 5 4 3 2 1 เสมอ ทำให้คะแนนที่ได้ในแต่ละข้ออยู่ในมาตรฐานอัตราภาค (Edward, 1957 อ้างถึงใน สุวิมล ติรภานันท์, 2551) และมีการกำหนดใช้ค่านี้อีก ๆ กันมา

3. ข้อความทั้งหมดต้องเป็นเรื่องเดียวกันหรือมีความเป็นเอกมิตี ในการใช้มาตรวัดแบบประเมินค่าจากคะแนนรวมของลิเคอร์ทซึ่งเป็นการรวมคะแนนเป็นรายข้อเข้าด้วยกัน ข้อคำถามทั้งหมดต้องวัดในเนื้อหาเดียวกัน และการที่ข้อคำถามทั้งหมดวัดในเนื้อหาเดียวกันนี้เองทำให้สามารถรวมคะแนนรายข้อเป็นคะแนนของมาตรวัดทั้งฉบับที่เป็นมาตรวัดอัตราภาค ดังนั้นควรมีการตรวจสอบความเป็นเอกมิตีก่อนรวมคะแนนทั้งฉบับ

4. ประเด็นที่ควรหลีกเลี่ยงในการสร้างข้อความ คือ หลีกเลี่ยงข้อความที่เป็นอดีตมากกว่าปัจจุบัน หลีกเลี่ยงข้อความที่เป็นจริงหรือสามารถตีความได้ว่าจริงเพราะหากข้อความที่ทำหน้าที่เป็นสิ่งเร้าเป็นเรื่องจริง ย่อมทำให้ผู้ตอบคำถามทุกคนตอบเหมือนกันหมด สิ่งที่ได้เป็นค่าคงที่ไม่ใช่ค่าของตัวแปรและไม่ใช้ความคิดเห็น หลีกเลี่ยงข้อความที่มีความหมายกำกวม ข้อความที่ไม่เกี่ยวข้องกับตัวแปรที่ศึกษาหรือไม่ต่อเนื่องกับข้อความอื่น ๆ ข้อความที่ทุกคนมีแนวโน้มตอบว่าเห็นด้วย โดยควรเลือกใช้ข้อความสั้น ๆ ที่เรียบง่าย ชัดเจน ตรงไปตรงมา เลี่ยงการใช้คำที่เป็นคำรวม เช่น ทั้งหมด เสมอ ๆ ไม่เคยเลย ไม่พบเลย ข้อความที่ซับซ้อน โดยในแต่ละคำถามต้องใช้ข้อความที่มีใจความเดียว เนื่องจากการใช้ข้อความที่มีใจความมากกว่า 1 ใจความทำให้ผู้ตอบสับสนว่าต้องการให้พิจารณาใจความส่วนใด โดยข้อสังเกตในการใช้มาตรฐานค่าแบบลิเคอร์ทมีดังนี้

4.1 การใช้มาตราประมาณค่าแบบลิเคอร์ท์ที่มีคำตอบ 5 ตัวเลือก ควรใช้ในลักษณะของผู้ตอบประเมินตนเอง ไม่เหมาะสมในการนำมาใช้สัมภาษณ์เพราะมีโอกาสเกิดความคลาดเคลื่อนได้มาก

4.2 ควรใช้กับผู้ตอบที่มีการศึกษามากพอสมควร เพื่อให้มีการตัดสินใจเลือกคำตอบที่ชัดเจน เพราะมักพบว่าผู้ตอบส่วนใหญ่มีการตอบตัวเลือกตรงกลาง 3 ตัวเลือกเท่านั้น คือ “เห็นด้วย” หรือ “ไม่เห็นด้วย” และ “ไม่เห็นด้วย”

4.3 ปัจจุบันเริ่มมีข้อถกเถียงเกี่ยวกับการใช้คะแนนรวมในการบอกทิศทางและระดับเจตคติของแต่ละบุคคล เพราะการที่ได้คะแนนรวมเท่ากันอาจมีเจตคติไม่เท่ากันก็ได้ ดังนั้นการใช้คะแนนรวมรายชื่อในมาตราชั้นบัญญัติอาจเป็นการนำเสนอผลที่ให้สารสนเทศรายชื่อชัดเจนมากกว่าคะแนนรวมจึงสามารถนำไปเป็นแนวทางการแก้ไขปรับปรุงในแต่ละประเด็น

2.2 ขั้นตอนในการพัฒนาเครื่องมือวัด

โดยมีขั้นตอนในการพัฒนาเครื่องมือวัด ดังนี้ (โชติกา ภาษีผล และคณะ, 2558; ศิริชัย กาญจนวาสี, 2556)

1. กำหนดจุดมุ่งหมายของการสอบ (specification of purpose)

1.1 วิเคราะห์จุดมุ่งหมายโดยการวิเคราะห์จุดมุ่งหมายของหลักสูตร จุดมุ่งหมายของการจัดการเรียนรู้ และระบุคุณลักษณะหรือสมรรถนะที่ต้องการวัดในการสอบ จากนั้นจึงแปลงคุณลักษณะหรือสมรรถนะที่ต้องการวัดให้มีลักษณะเป็นรูปธรรม โดยเขียนเป็นจุดประสงค์เชิงพฤติกรรมตามระดับพฤติกรรมที่เป็นจุดมุ่งหมายที่ต้องการให้เกิดกับผู้เรียนเมื่อเรียนจบวิชานั้นแล้ว

1.2 วิเคราะห์เนื้อเรื่อง โดยการวิเคราะห์เนื้อหาของหลักสูตร เนื้อหาที่จัดการเรียนรู้เพื่อแยกแยะเนื้อหา และจัดรวมเป็นหน่วย เพื่อเรียงลำดับความสัมพันธ์ของเนื้อหาตามลำดับก่อน-หลัง

1.3 วิเคราะห์กิจกรรมหรือประสบการณ์ เพื่อกำหนดรูปแบบวิธีการจัดการเรียนรู้และวิธีสอบที่เหมาะสมกับจุดมุ่งหมายและเนื้อหาของการเรียนรู้

2. ออกแบบการสร้างแบบสอบ (test design) เป็นการกำหนดรูปแบบ ขอบเขต และแนวทางการสร้างเพื่อให้ได้มาซึ่งข้อสอบและแบบสอบที่มีคุณภาพมีขั้นตอน ดังนี้

2.1 วางแผนการทดสอบ (testing plans) ผู้สอนจะต้องมีการวางแผนการสอบให้สอดคล้องกับแผนการวัดผลที่กำหนดไว้ อาจเป็นการวัดผลก่อนการจัดการเรียนรู้ (placement testing) เพื่อดูความพร้อมหรือจัดกลุ่มผู้เรียนก่อนเริ่มเรียน การทดสอบความก้าวหน้าระหว่างการจัดการเรียนรู้ (formative testing) หรือการทดสอบเพื่อวินิจฉัยสาเหตุที่ทำให้นักเรียนเกิดความยุ่งยากในการเรียนรู้ (diagnostic testing) ระหว่างการจัดการเรียนรู้ หรือมีการสอบหลังเสร็จสิ้นการจัดการเรียนรู้หรือการทดสอบสรุปรวม (summative testing)

2.2 กำหนดรูปแบบของแบบสอบ (test format) ผู้สอบจะต้องเลือกรูปแบบของแบบสอบว่ารูปแบบสอบใดที่น่าจะเหมาะสมกับสมรรถภาพและเนื้อหาที่มุ่งวัดสำหรับการทดสอบแต่ละครั้ง

2.3 สร้างผังข้อสอบ (test blueprint) หรือ ตารางกำหนดแผนผังการสร้างข้อสอบ (table of specification) เป็นตารางที่สร้างขึ้นเพื่อนำเสนอรายละเอียดของการทดสอบแต่ละครั้งว่าจะวัดเนื้อหาใด มีจุดประสงค์ของการเรียนรู้หรือวัตถุประสงค์เชิงพฤติกรรมคืออะไร อาจมีน้ำหนักความสำคัญหรือสัดส่วนข้อสอบที่จะสร้างขึ้นเป็นแบบสอบ

3. เขียนข้อสอบ (item writing)

3.1 กำหนดแบบแผนผังข้อสอบ (item specification) ผู้สอนควรกำหนดแบบแผนผังข้อสอบซึ่งจะช่วยประหยัดเวลาในการเขียนข้อสอบในครั้งต่อไป สามารถสร้างเป็นข้อสอบได้หลายฉบับ และพัฒนาเป็นแบบสอบคู่ขนานได้

3.2 ร่างข้อสอบ (item drafting) ผู้สอบลงมือร่างข้อสอบตามแบบแผนข้อสอบแต่ละจุดประสงค์ให้ครอบคลุมจุดประสงค์เชิงพฤติกรรมและเนื้อหาที่ต้องการวัด การร่างข้อสอบควรเผื่อข้อสอบไว้ 25% หรือหากเป็นไปได้ควรเผื่อไว้ประมาณ 1-2 เท่าของจำนวนที่ต้องการใช้จริง เพื่อไว้สำหรับตัดออกหรือปรับปรุงข้อที่ไม่ดีและเลือกข้อที่มีคุณภาพดีไว้ใช้ต่อไป

3.3 ทบทวนร่างข้อสอบ (item review) แบ่งเป็นการทบทวนร่างข้อสอบโดยผู้เขียนข้อสอบ และการทบทวนร่างข้อสอบโดยผู้เชี่ยวชาญเพื่อช่วยตรวจสอบร่างข้อสอบโดยพิจารณาถึงความตรงของข้อสอบตามจุดประสงค์มีความครอบคลุม ความชัดเจนของทั้งคำถามและคำตอบ

ในการทบทวนร่างข้อสอบโดยผู้เชี่ยวชาญควรผู้สร้างควรให้ผู้เชี่ยวชาญทางด้านเนื้อหา และการวัดประเมินผลตรวจสอบความเหมาะสมของนิยามสิ่งที่มีวัด ความครอบคลุมของเนื้อหาและสัดส่วนของคำถามที่ใช้ และตรวจสอบความสอดคล้องระหว่างข้อคำถามกับเนื้อหาหรือจุดประสงค์ที่มีวัด โดยปกติจะให้ผู้เชี่ยวชาญจำนวนไม่น้อยกว่า 3 คนขึ้นไป เป็นผู้พิจารณา ผู้เชี่ยวชาญแต่ละท่านจะได้รับ (1) ผังข้อสอบ (test blueprint) (2) ข้อสอบ และ (3) แบบฟอร์มการตรวจ

การให้ผู้เชี่ยวชาญพิจารณาความสอดคล้องระหว่างข้อคำถามกับเนื้อหาหรือจุดประสงค์ (ความตรงตามเนื้อหา) หรือในกรณีที่สร้างเครื่องมือวัดคุณลักษณะอื่นๆ จะเป็นการพิจารณาความสอดคล้องกับทฤษฎีที่อธิบายคุณลักษณะนั้น (ความตรงตามโครงสร้าง) โดยสามารถใช้สูตรการคำนวณดัชนี Item Objective Congruence หรือ IOC ดังนี้ (ณัฐภรณ์ หลาวทอง, 2551)

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

เมื่อ

IOC	คือ	ดัชนีความสอดคล้องระหว่างข้อคำถามกับเนื้อหา/วัตถุประสงค์/ทฤษฎี
R	คือ	คะแนนผลการตัดสินข้อคำถามของผู้เชี่ยวชาญ
	+1	หมายถึง แน่ใจ ว่าข้อคำถามวัดได้ตรงกับเนื้อหา/วัตถุประสงค์/ทฤษฎี
	0	หมายถึง ไม่แน่ใจ ว่าข้อคำถามวัดได้ตรงกับเนื้อหา/วัตถุประสงค์/ทฤษฎี
	-1	หมายถึง แน่ใจ ว่าข้อคำถามวัดได้ ไม่ตรง กับเนื้อหา/วัตถุประสงค์/ทฤษฎี
N	คือ	จำนวนผู้เชี่ยวชาญ

เกณฑ์ในการแปลความหมาย

ถ้า $IOC \geq 0.5$ แสดงว่าข้อสอบนั้นวัดได้สอดคล้องกับเนื้อหา/วัตถุประสงค์/ทฤษฎี

ถ้า $IOC < 0.5$ แสดงว่าข้อสอบนั้นวัดไม่สอดคล้องกับเนื้อหา/วัตถุประสงค์/ทฤษฎี

3.4 บรรณาธิการข้อสอบ (item editing) เป็นการปรับปรุงข้อบกพร่องตามคำแนะนำที่ได้รับจากผู้เชี่ยวชาญ ปรับภาษาที่ใช้ให้เหมาะสมกับผู้เรียน เรียบเรียงข้อสอบรวมเป็นแบบสอบพร้อมที่จะนำไปทดลองใช้

4. ทดลองใช้ข้อสอบและวิเคราะห์ข้อสอบ (item tryout and analysis)

หลังจากที่ได้สร้างแบบสอบและปรับปรุงคุณภาพตามคำแนะนำของผู้เชี่ยวชาญเป็นที่เรียบร้อยแล้ว ขั้นตอนต่อไปคือการนำแบบสอบที่สร้างขึ้นไปทดลองใช้กับกลุ่มที่มีความใกล้เคียงกับกลุ่มเป้าหมายที่ต้องการนำไปใช้จริง และนำผลที่ได้จากการทดลองใช้มาวิเคราะห์ข้อสอบดังนี้

4.1 การวิเคราะห์ข้อสอบ แบ่งเป็น

(1) การวิเคราะห์ทางกายภาพ ทั้งด้านความชัดเจนของคำสั่งคำถาม คำตอบ ความเหมาะสมของภาษา ความยาวของแบบสอบ ระยะเวลาที่กำหนดให้ รูปแบบการพิมพ์ เป็นต้น

(2) การวิเคราะห์เชิงปริมาณ หรือการวิเคราะห์ข้อสอบ (item analysis) เพื่อให้ทราบคุณภาพของข้อสอบเป็นรายข้อเกี่ยวกับ ความยากง่ายข้อสอบ (item difficulty) อำนาจจำแนกของข้อสอบ (item discrimination) และประสิทธิภาพของตัวลวง (effectiveness of distracters)

4.2 การคัดเลือกข้อสอบรวมเป็นแบบสอบ (assembling the test) การคัดเลือกข้อสอบที่มีความเหมาะสมที่พิจารณาทั้งความยากง่ายข้อสอบ อำนาจจำแนกของข้อสอบ รวมถึงประสิทธิภาพของตัวลวงของข้อสอบรายข้อเพื่อใช้เป็นเกณฑ์ในการคัดเลือกข้อสอบที่ดีและครอบคลุมเนื้อหาที่ต้องการวัด เพื่อนำมาจัดเรียงเป็นแบบสอบฉบับที่จะนำมาใช้ต่อไป

4.3 การวิเคราะห์แบบสอบ (test analysis) เมื่อคัดเลือกข้อสอบที่ดีที่มีคุณภาพเป็นรายข้อมารวมกันเป็นแบบสอบแล้ว ควรมีการวิเคราะห์แบบสอบที่ได้ทั้งด้านความเที่ยง (reliability) และความตรง (validity)

5. นำแบบสอบไปใช้ การนำแบบสอบไปใช้ควรคำนึงถึงปัจจัยต่างๆที่มีอิทธิพลต่อการแสดงความสามารถในการตอบคำถามของผู้เรียน ดังนี้

5.1 คำชี้แจง (directions) โดยปกติจะต้องประกอบด้วยคำสั่ง 2 ชุด สำหรับผู้คุม (test administrator) และผู้สอบ (test taker) ถ้าเป็นแบบสอบมาตรฐานคำสั่งของผู้คุมสอบจะต้องพิมพ์แยกออกมาต่างหากจากตัวแบบสอบ คำสั่งจะให้แนวทาง/ทิศทางที่ชัดเจนในการจัดห้องสอบ แจกข้อสอบ เก็บข้อสอบ เวลาที่ใช้ และแนวทางแก้ปัญหาที่อาจเกิดขึ้นระหว่างการสอบซึ่งผู้คุมจะต้องถือปฏิบัติอย่างเคร่งครัด

5.2 กำหนดเวลาของการตอบ (time limits) เวลาที่ใช้ขึ้นอยู่กับประเภทของข้อสอบ ความซับซ้อนของจุดประสงค์การเรียนรู้ที่ต้องการวัด ระดับอายุของผู้ตอบ หากเป็นข้อสอบแบบ power test ควรใช้เวลาให้ผู้สอบส่วนใหญ่ ประมาณร้อยละ 90 สามารถทำข้อสอบได้ครบถ้วน แต่ทั้งนี้ในทางปฏิบัติเวลาที่ใช้ในการสอบจะมีความสัมพันธ์กับสัดส่วนของจำนวนชั่วโมงเรียนหรือจำนวนหน่วยกิต ดังนั้น ผู้สอบจึงต้องเลือกประเภทของข้อสอบ จำนวนข้อ และระดับความซับซ้อนของจุดประสงค์การเรียนรู้ที่มุ่งวัดให้เหมาะสมกับเวลาและผู้สอบ

5.3 เงื่อนไขในการสอบ (testing conditions) ต้องมีการบริหารจัดการการสอบให้เหมาะสมและเป็นมาตรฐานเดียวกันทั้งสิ่งแวดล้อมทางกายภาพ (physical environment) ได้แก่ สภาพแวดล้อมทั่วไป เช่น ห้องสอบมีความกว้างขวาง และแสงสว่างเพียงพอ ปราศจากเสียงหรือกลิ่นรบกวน มีอุณหภูมิที่เหมาะสม อากาศถ่ายเทได้ดี และสิ่งแวดล้อมทางจิตวิทยา (psychological environment) โดยเฉพาะ "ความเครียดและความวิตกกังวล" ที่อาจเป็นสาเหตุให้ผู้สอบไม่อาจแสดงความสามารถสูงสุดออกมาหากมีความเครียดมากเกินไป ดังนั้น ก่อนการสอบผู้สอนควรบอกให้ผู้สอบรู้ตัวล่วงหน้า รู้จุดประสงค์ของการสอบ ขอบเขตเนื้อหาที่สอบนี้หนักความสำคัญของการสอบ และการนำผลสอบไปใช้ ในขณะระหว่างการสอบ ผู้คุมสอบจะต้องไม่สร้างบรรยากาศที่ก่อให้เกิดความเครียด หรือรบกวนในการสอบ

5.4 การตรวจให้คะแนน (scoring) การตรวจให้คะแนนต้องมีความเป็นปรนัย ซึ่งสามารถทำได้โดย ควรมีการบันทึกคำตอบที่ชัดเจนและสมบูรณ์ มีการเตรียมคำตอบถูกไว้สำหรับตรวจให้คะแนน (scoring key) และมีการระบุเกณฑ์การให้คะแนน (scoring criteria)

5.5 การนำผลไปใช้ (using the result) รายงานผลการสอบ และนำผลการสอบไปใช้ประโยชน์ในการพัฒนาการเรียนรูของผู้เรียนและปรับปรุงการเรียนการสอนของผู้สอน

6. วิเคราะห์คุณภาพของแบบสอบ (test analysis) หลังจากที่ได้นำแบบสอบไปใช้แล้วควรมีการดำเนินการดังนี้

6.1 การวิเคราะห์ค่าสถิติเบื้องต้นของการทดสอบ การวิเคราะห์ด้วยค่าสถิติพื้นฐานได้แก่ ค่าเฉลี่ย (Mean: M) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation: SD) รวมถึงการหาคะแนนสูงสุดและต่ำสุดของการทดสอบนั้น นอกจากนี้หากเป็นไปได้ควรมีการวิเคราะห์รูปทรงและการแจกแจงของคะแนนสอบ ซึ่งจะทำให้ได้สารสนเทศที่เป็นประโยชน์ต่อการแปลความหมายของคะแนนแบบสอบและการปรับปรุงแบบสอบต่อไป โดยเฉพาะแบบสอบอิงกลุ่ม

6.2 การวิเคราะห์แบบสอบ ควรมีการวิเคราะห์คุณภาพของแบบสอบที่นำไปใช้จริงอีกครั้งทั้งค่าความยาก อำนาจจำแนก และความเที่ยงของแบบสอบทั้งฉบับ รวมทั้งการหาความตรงบางชนิด เช่น ความตรงตามเกณฑ์สัมพัทธ์ หรือความตรงเชิงโครงสร้าง

7. ปรับปรุงแบบสอบ (test revision) มีการปรับปรุงแบบสอบตามข้อบกพร่องที่พบเพื่อนำไปใช้กับกลุ่มอื่นๆ ที่มาจากประชากรเป้าหมายเดียวกัน การนำไปใช้ควรเป็นไปตามเงื่อนไขหรือมาตรฐานตามที่กำหนดไว้ แล้วทำการวิเคราะห์ซ้ำอีก ถ้าผลการวิเคราะห์ยืนยันว่าเป็นแบบสอบที่มีคุณภาพควรมีการพัฒนาปกติวิสัย (norm) หรือเกณฑ์ เพื่อเป็นบรรทัดฐานของการเปรียบเทียบความหมายของคะแนน และเก็บไว้ในคลังข้อสอบต่อไป

2.3 การตรวจสอบคุณภาพเครื่องมือ

การวิเคราะห์ตรวจสอบคุณภาพของเครื่องมือได้มีผู้ให้แนวคิดและวิธีการหาคุณภาพ สรุปได้ดังนี้

1. คุณภาพด้านความตรง (validity)

บุญธรรม กิจปริดาภิสุทธิ์ (2553) กล่าวว่า เครื่องมือเก็บรวบรวมข้อมูลที่ตี ต้องมีความตรง (validity) ในการวัด หมายความว่า วัดได้ตรงตามวัตถุประสงค์และพฤติกรรมที่ต้องการให้วัด วัดได้ครอบคลุมครบถ้วนตามเนื้อหาที่ต้องการให้วัด และวัดได้ถูกต้องตรงตามความเป็นจริง ความตรงของเครื่องมือเก็บรวบรวมข้อมูลแบ่งได้ 3 ประเภท คือ

1) ความตรงตามเนื้อหา (content validity) หมายถึง วัดได้ตรงตามเนื้อหา และครอบคลุมครบถ้วนตามเนื้อหาที่ต้องการให้วัด

2) วัดตรงตามเกณฑ์ (criterion-related validity) หมายถึง วัดได้ตรงหรือเหมือน กับเกณฑ์ที่ต้องการให้วัด ถ้าตรงหรือเหมือนกับสภาพความเป็นจริงในปัจจุบัน เรียกว่า ความตรงตามสภาพการณ์ (concurrent validity) ถ้าวัดได้ตรงหรือเหมือนกับสภาพความเป็นจริงในอนาคต เรียกว่า ความตรงเชิงพยากรณ์ (predictive validity)

3) ความตรงตามโครงสร้าง (construct validity) หมายถึง วัดได้ตรงตามพฤติกรรมตามทฤษฎีหรือแนวคิดของเครื่องมือที่ต้องการให้วัด

เครื่องมือรวบรวมข้อมูลในการวิจัยนั้น ต้องความตรงโดยเฉพาะความตรงตามเนื้อหา และความตรงตามโครงสร้างการจะสร้างให้มีความตรงดังกล่าว ก่อนสร้างจะต้องกำหนดเนื้อหาและพฤติกรรมให้ชัดเจนว่า ต้องการสร้างให้วัดเนื้อหาและพฤติกรรมอะไรบ้าง ในปริมาณมากน้อยเพียงใด เมื่อสร้างแล้วต้องมีการตรวจสอบโดยอาศัยความเห็นของผู้เชี่ยวชาญและการนำไปทดลองใช้

ศิริชัย กาญจนวาสี (2556) ได้กล่าวว่า ความตรง (validity) เป็นคุณสมบัติที่สำคัญของเครื่องมือวัดผล ซึ่งเกี่ยวกับคุณภาพด้าน ความถูกต้อง ของผลที่ได้จากการวัด ทำให้สามารถนำคะแนนที่ได้ไปแปลความหมาย ถึงสิ่งที่มุ่งวัดได้อย่างเหมาะสม การตรวจสอบความตรงสามารถจำแนกตามเป้าหมายสำคัญได้ 3 ประเภทคือ

1. การตรวจสอบความตรงตามเนื้อหา (content validation) เป็นการสรุปอ้างอิงถึง มวลเนื้อเรื่อง ความรู้ หรือประสบการณ์ (domain of content) ที่แบบสอบมุ่งวัดว่า การวัดของแบบสอบได้ผลครอบคลุม และเป็นตัวแทนมวลความรู้หรือประสบการณ์ที่มุ่งวัดเพียงใด

2. การตรวจสอบความตรงตามเกณฑ์สัมพัทธ์ (criterion-related validation) เป็นการสรุปอ้างอิงถึง สมรรถนะ การดำเนินงาน (practical performance) ของสิ่งที่มุ่งวัดว่า การวัดของแบบทดสอบได้ผลสอดคล้องกับการดำเนินงานของคุณลักษณะนั้นเพียงไร ซึ่งเกณฑ์ที่ใช้เทียบความสัมพันธ์อาจเป็น สถานภาพการดำเนินงานที่เป็นอยู่จริงในปัจจุบัน หรืออาจเป็นผลสำเร็จของการปฏิบัติงานนั้นในอนาคต

3. การตรวจสอบความตรงเชิงโครงสร้าง (construct validation) เป็นการสรุปอ้างอิงถึง สภาวะ หรือ โครงสร้างทางจิตวิทยา (psychological construct) ของลักษณะที่มุ่งวัดนั้นว่า การวัดได้ผลตรงตามทฤษฎีของลักษณะทางจิตวิทยาของคุณลักษณะนั้นดีเพียงไร อย่างไรก็ดี ศิริชัย กาญจนวาสี (2556) กล่าวว่า การตรวจสอบความตรงเชิงทฤษฎีจำเป็นต้องอาศัยการสนับสนุนอ้างอิงถึงลักษณะที่มุ่งวัด โดยใช้หลักฐานสนับสนุนหลายอย่าง ความตรงตามเนื้อเรื่องและความตรงตามจุดสำคัญถือว่าเป็นหลักฐานส่วนหนึ่งสำหรับสนับสนุนความตรงเชิงทฤษฎี สำหรับสรุปผลความตรงเชิงทฤษฎีจำเป็นต้องอาศัยหลักฐานความสอดคล้องกับทฤษฎี เช่น หลักฐานการตรวจสอบผลการวัดที่เป็นไปตามคำทำนายของทฤษฎี หลักฐานการเปรียบเทียบความสัมพันธ์กับผลการวัดด้วยเครื่องมือวัดที่มุ่งวัดลักษณะเดียวกัน และลักษณะที่ต่างกัน เป็นต้น มีวิธีการสำหรับหาหลักฐานเชิงประจักษ์เพื่อใช้สนับสนุนความตรงเชิงทฤษฎีหลายวิธีวิธีที่นิยมใช้กันมีดังนี้

3.1 วิธีตัดสินโดยผู้เชี่ยวชาญ หลักฐานเบื้องต้นอย่างหนึ่งที่สามารถนำมาใช้สนับสนุนความตรงเชิงทฤษฎีของ แบบสอบ ได้แก่ การตัดสินของกลุ่มผู้เชี่ยวชาญ โดยให้กลุ่มผู้เชี่ยวชาญตรวจสอบความเหมาะสมของทฤษฎีที่นำมาใช้ นิยาย ผังข้อสอบ และคุณภาพของข้อสอบ

3.2 วิธีเปรียบเทียบคะแนนระหว่างกลุ่มที่ทราบ (comparing the scores of known groups) ในกรณีที่มีความเที่ยงตามทฤษฎีว่า คะแนนผลการวัดลักษณะที่สนใจนั้นมีความแตกต่างกันระหว่างกลุ่มผู้สอบที่ทราบแน่ชัดว่ามีลักษณะสำคัญบางประการแตกต่างกัน เช่น แตกต่างกันตามกลุ่มอายุ กลุ่มความสามารถกลุ่มที่ผ่านประสบการณ์เฉพาะอย่าง กลุ่มที่มีการปรับตัวได้ดี/ปรับตัวไม่ดี กลุ่มที่ประสบความสำเร็จ/ไม่ประสบความสำเร็จ เป็นต้น จึงน่าจะเป็นหลักฐานส่วนหนึ่งที่สนับสนุนความตรงเชิงทฤษฎีด้วย การเปรียบเทียบคะแนนระหว่างกลุ่มที่ทราบแน่ชัดแล้วว่ามีลักษณะที่มุ่งวัดนั้นแตกต่างกัน (known groups) ถ้าเครื่องมือสามารถวัดลักษณะที่สนใจนั้นได้ ผลการวัดจะต้องมีความแตกต่างระหว่างกลุ่ม

3.3 วิธีวิเคราะห์ตัวประกอบ (factor analysis) หลักฐานอย่างหนึ่งที่สามารถนำมาใช้สนับสนุนความตรงเชิงทฤษฎีได้แก่ ความตรงตามตัวประกอบ (factorial validity) ซึ่งเป็นคุณสมบัติของการประกอบที่มุ่งวัด สามารถศึกษาได้ด้วยการวิเคราะห์องค์ประกอบ (factor analysis) ซึ่งเป็นเทคนิคทางสถิติสำหรับวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรที่สังเกตค่าได้เพื่อหาลักษณะร่วมกันของชุดตัวแปรเหล่านั้น ลักษณะร่วมกันนี้เรียกว่า ตัวประกอบ (factor) เป็นลักษณะที่คาดว่ามียุทธิพลของคะแนนที่ได้จากกลุ่มตัวแปร หรือเป็นลักษณะที่ใช้อธิบายความผันแปรของกลุ่มตัวแปร ทั้งนี้ โมเดลสำหรับวิเคราะห์ตัว

ประกอบที่ใช้อยู่ในปัจจุบันมีอยู่ 2 โมเดล ได้แก่โมเดลการวิเคราะห์ตัวประกอบเชิงสำรวจ (Exploratory Factor Model: EFA) และโมเดลการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Model: CFA)

2. คุณภาพด้านความเที่ยง (Reliability)

ศิริชัย กาญจนวาสี (2556) สรุปนิยามเชิงทฤษฎีของความเที่ยงว่า ความเที่ยง หมายถึง ความคงที่หรือความคงเส้นคงวาของผลที่ได้จากการวัดซ้ำ ถ้าการวัดสิ่งเดียวกันหลาย ๆ ครั้งได้ค่าคงที่ค่อนข้างคงเส้นคงวาสูงขึ้นเพียงใดก็ถือว่าการวัดมีความเที่ยงมากขึ้นเพียงนั้น แบบสอบที่มีความเที่ยงสูงจะเป็นเครื่องมือที่ใช้วัดคุณลักษณะที่ต้องการได้ผลอย่างคงเส้นคงวา ปัญหาความเที่ยงของแบบสอบเกี่ยวข้องกับคำถามที่ว่า แบบสอบสามารถให้ผลการวัดที่คงเส้นคงวาเพียงใด ถ้าทำการวัดซ้ำ ๆ ส่วนในทางปฏิบัติสามารถให้นิยามเชิงปฏิบัติการของความเที่ยงของแบบสอบ หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบสอบคู่ขนาน 2 ชุด ซึ่งสอบโดยผู้สอบกลุ่มเดียวกัน ด้วยวิธีดำเนินการสอบที่เป็นอิสระจากกันแต่เป็นมาตรฐานเดียวกัน อย่างไรก็ตามอย่างไรก็ดี ศิริชัย กาญจนวาสี (2556) ได้กล่าวถึงวิธีการประมาณค่าความเที่ยงของเครื่องมือทำได้หลายวิธี ดังนี้

2.1 วิธีการประมาณค่าความเที่ยงแบบอิงเกณฑ์

1) วิธีสอบซ้ำ (test-retest method)

การประมาณค่าความเที่ยงโดยวิธีสอบซ้ำกระทำได้โดยนำแบบสอบฉบับเดียวกันไปทดสอบ ผู้สอบกลุ่มเดียวกัน 2 ครั้ง โดยทิ้งช่วงระยะเวลาระหว่างการสอบครั้งแรกและครั้งที่สอง แล้วนำคะแนนที่ได้จากการสอบครั้งแรก กับครั้งหลัง มาคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (pearson product moment correlation coefficient)

2) วิธีใช้แบบสอบสมมูล (equivalent-forms method)

การประมาณค่าความเที่ยงโดยวิธีนี้ สามารถดำเนินการได้โดยนำแบบสอบ 2 ฉบับที่สมมูลกัน (equivalent-forms or alternate-forms) ไปทดสอบกับผู้สอบกลุ่มหนึ่งในเวลาเดียวกัน แล้วนำคะแนนที่ได้จากแบบสอบทั้งสองฉบับมาคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

3) วิธีสอบซ้ำด้วยแบบสอบสมมูล (test-retest with equivalent-forms)

การประมาณค่าความเที่ยงโดยวิธีสอบซ้ำด้วยแบบสอบสมมูล เป็นวิธีดำเนินงาน ที่ผสมวิธีการสอบซ้ำ และวิธีใช้แบบสอบสมมูลเข้าด้วยกัน โดยทำการสอบผู้สอบกลุ่มเดียวกัน 2 ครั้ง ที่ต่างเวลากันด้วยแบบสอบ 2 ฉบับที่สมมูลกัน แล้วนำคะแนนที่ได้จากแบบสอบทั้งสองฉบับมาคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

4) วิธีตรวจสอบความสอดคล้องภายใน (internal consistency method)

การตรวจสอบความสอดคล้องภายในของแบบสอบเป็นวิธีการประมาณค่าความเที่ยงของแบบสอบโดยใช้การทดสอบเพียงครั้งเดียว ด้วยการใช้แบบสอบฉบับเดียวทำการทดสอบผู้สอบกลุ่มเดียวกัน ในการตรวจสอบความสอดคล้องภายในนั้นเป็นการวัดระดับความเป็นเอกพันธ์ (homogeneity) ของข้อสอบในแบบสอบนั้นว่า วัดเนื้อเรื่องเดียวกันเพียงใด ถ้าแบบสอบวัดในเรื่องเดียวกัน เมื่อทำการวัดซ้ำๆ ก็น่าจะมีค่าคงที่หรือสอดคล้องในผลการวัดสูง สำหรับวิธีหาความเที่ยงแบบสอดคล้องภายในที่นิยมใช้กันมีดังนี้

4.1) วิธีแบ่งครึ่งข้อสอบ (split-half methods)

การประมาณค่าความเที่ยงแบบความสอดคล้องภายในโดยวิธีแบ่งครึ่งข้อสอบสามารถกระทำได้โดยนำแบบสอบที่ต้องการหาความเที่ยงไปทดสอบกับกลุ่มผู้สอบ เมื่อตรวจคะแนนเป็นรายข้อแล้ว แบ่งการรวมคะแนนของข้อสอบออกเป็น 2 ส่วน โดยให้กลุ่มข้อสอบทั้ง 2 ส่วนสมดุลกันมากที่สุด จากนั้นนำคะแนนสองส่วนดังกล่าวมาคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ซึ่งจะได้ค่าสัมประสิทธิ์ความเที่ยงเพียงครึ่งฉบับ ดังนั้นจึงต้องใช้สูตรของสเปียร์แมนบราวน์ (Spearman-Brown formula) มาปรับให้เป็นสัมประสิทธิ์ความเที่ยงเต็มฉบับ

4.2) วิธีสัมประสิทธิ์แอลฟาของครอนบาค (Cronbach's Alpha method)

เทคนิคของการแบ่งครึ่งข้อสอบเพื่อนำคะแนนทั้งสองส่วนมาคำนวณความสัมพันธ์สำหรับประมาณค่าความเที่ยงแบบความสอดคล้องภายในของแบบสอบนั้น สามารถขยายแนวคิดไปใช้กับการแบ่งแบบสอบออกเป็นมากกว่า 2 ส่วน หรือ k ส่วน เมื่อคำนวณความแปรปรวนของคะแนนแต่ละส่วนและความแปรปรวนของคะแนนรวม เราสามารถนำไปใช้ประมาณค่าความเที่ยงแบบความสอดคล้องภายในได้

4.3) วิธีของคูเดอร์-ริชาร์ดสัน (Kuder-Richardson's method)

การหาความสอดคล้องกันระหว่างข้อคำถามในเครื่องมือชุดเดียวกัน ดำเนินการใช้ครั้งเดียว สำหรับแบบสอบที่มีการให้คะแนนแบบ 0 หรือ 1 เท่านั้น การคำนวณค่าดัชนีความเที่ยงโดยใช้สูตร KR20 หรือ KR21 นั้นใช้ในกรณีที่ข้อสอบปรนัย คือถูกได้หนึ่ง ผิดได้ศูนย์ สูตร KR21 นั้น เป็นสูตรที่ง่ายต่อการคำนวณแต่มีข้อจำกัดที่ว่าค่าดัชนีความเที่ยง ที่ได้จากการคำนวณโดยใช้สูตร KR21 นั้นจะต่ำกว่าที่ควรจะเป็น ถ้าแบบทดสอบประกอบด้วยข้อสอบที่มีดัชนีความง่ายแตกต่างกัน

2.2 วิธีการประมาณค่าความเที่ยงแบบอิงเกณฑ์

1) วิธีของคาร์เวอร์ (Carver method)

วิธีของคาร์เวอร์เป็นแนวคิดที่ได้จากการตรวจสอบความสอดคล้องของผู้รอบรู้และไม่รอบรู้ ซึ่งสามารถทำได้ 2 แนวทางคือ แนวทางที่ 1 ให้นักเรียนทำแบบสอบคู่ขนาน 2 ฉบับ ในการทดสอบครั้งเดียวกัน และแนวทางที่ 2 ใช้แบบสอบ 1 ฉบับ ทำการทดสอบซ้ำกับนักเรียนกลุ่มเดียวกัน โดยทั้งสองระยะเวลาให้เหมาะสม และกำหนดจุดตัดคะแนนระหว่างผู้รอบรู้และไม่รอบรู้ (ศิริชัย กาญจนวาสี, 2556; สมนึก ภัททิยธนี, 2560)

2) วิธีของโลเวทท์ (Lovett method)

วิธีของโลเวทท์เป็นการนำแบบสอบแบบอิงเกณฑ์ 1 ฉบับ ไปทดสอบผู้เรียน 1 กลุ่มเพียงครั้งเดียว แล้วจึงนำผลที่ได้มาคำนวณหาค่าความเที่ยง

2.3 ความสอดคล้องระหว่างผู้ประเมิน (inter-rater reliability)

ความสอดคล้องระหว่างผู้ประเมิน (inter-rater reliability หรือ inter-rater agreement) เป็นดัชนีค่าความเที่ยงที่มีความสำคัญในกรณีที่ผู้สอนใช้เครื่องมือประเภทข้อสอบอัตนัยหรือข้อสอบความเรียง การประเมินทักษะการปฏิบัติ (performance assessment) และการประเมินด้วยแฟ้มสะสมผลงาน (portfolio assessment) ซึ่งจำเป็นต้องมีการพัฒนาและใช้เกณฑ์การให้คะแนนแบบรูบริก (scoring rubric) ประกอบด้วย ดังนั้น เพื่อยืนยันความสอดคล้องของผลการประเมินระหว่างผู้ประเมินตั้งแต่ 2 คนขึ้นไป โดยคำนวณความสัดส่วนของผู้รับการประเมินที่ผู้ประเมิน 2 คนให้คะแนนตรงกัน ต่อจำนวนผู้รับการประเมินทั้งหมด คูณด้วย 100 (กมลวรรณ ตังธนาภานนท์, 2557; Nitko & Brookhart, 2007; Reynolds, et al., 2009)

3. คุณภาพด้านความยาก (difficulty)

ความยาก หมายถึง ร้อยละหรือสัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นได้ถูกต้อง มีค่าอยู่ระหว่าง 0 ถึง 1 กล่าวคือ หากค่าเข้าใกล้ 0 แสดงว่าข้อสอบข้อนั้นยาก แต่หากค่าเข้าใกล้ 1 แสดงว่าข้อสอบข้อนั้นง่าย การวิเคราะห์คุณภาพด้านความยากนี้ จะพิจารณาในเครื่องมือที่เป็นแบบทดสอบเท่านั้น

4. คุณภาพด้านอำนาจจำแนก (discrimination)

อำนาจจำแนก หมายถึง ความสามารถของข้อสอบหรือข้อคำถามในการจำแนกหรือแยกให้เห็นความแตกต่างระหว่างผู้ทดสอบที่มีคุณสมบัติแตกต่างกันออกจากกันได้ อำนาจจำแนกหาได้จากอัตราส่วนของคนตอบถูกในกลุ่มสูงลบด้วยอัตราส่วนของคนตอบถูกในกลุ่มต่ำ ซึ่งมีค่าอยู่ระหว่าง -1 ถึง 1 แต่ค่าอำนาจจำแนกที่ดีต้องมีค่าเป็นบวก โดยมีค่าตั้งแต่ 0.2 ขึ้นไป

กรณีเครื่องมือไม่ใช่แบบทดสอบ สามารถหาค่าอำนาจจำแนก ได้จากสูตรสัมประสิทธิ์สหสัมพันธ์แบบ Pearson Product Moment (R_{xy}) สหสัมพันธ์แบบ Point-biserial (r_{pbis}) และสหสัมพันธ์แบบ biserial (r_{bis}) ซึ่งเป็นการหาค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม ถ้าค่าสัมประสิทธิ์สหสัมพันธ์ของข้อสอบข้อใดมีค่าเป็นบวกสูง คือ เข้าใกล้ +1 แสดงว่าผู้สอบที่เก่งหรือได้คะแนนสูงจากแบบสอบมีแนวโน้มที่จะตอบข้อสอบข้อนั้นถูก ในขณะที่ผู้สอบที่อ่อนหรือได้คะแนนต่ำจากแบบสอบมีแนวโน้มที่จะตอบข้อสอบข้อนั้นผิด จึงแปลความได้ว่าข้อสอบข้อนั้นมีอำนาจจำแนกสูง แต่หากค่าสัมประสิทธิ์สหสัมพันธ์มีค่าติดลบจะแปลความหมายในทิศทางตรงกันข้าม (โชติกา ภาชีผล 2554; ศิริชัย กาญจนवासี, 2556; สุวิมล ติรภานันท์, 2551)

2.4 การสร้างเกณฑ์ปกติ (norms)

ลิวัน สายยศและอังคณา สายยศ (2539) อ้างอิงมาจาก สมนึก ภัททิยธนี, 2560) ได้ให้ความหมายของเกณฑ์ปกติ (norms) หมายถึง ข้อเท็จจริงทางสถิติการแจกแจงของคะแนนประชากรที่นิยามไว้อย่างดีแล้ว เป็นคะแนนตัวที่จะบอกระดับความสามารถของผู้สอบว่าอยู่ระดับใดของกลุ่มประชากร การสร้างเกณฑ์ปกติคำนึงถึงหลัก 3 ประการดังนี้

1. ความเป็นตัวแทนที่ดี หมายถึง กลุ่มตัวอย่างที่นำมาใช้สร้างเกณฑ์ปกติ ต้องเกิดจากการสุ่มที่เป็นตัวแทนที่ดีของประชากร การสุ่มกลุ่มตัวอย่างโดยอาศัยความน่าจะเป็น ทำได้หลายวิธีทั้งนี้ต้องเลือกวิธีสุ่มให้เหมาะสมกับลักษณะประชากร

2. มีความเที่ยงตรง (validity) หมายถึง การนำคะแนนสอบไปเทียบกับเกณฑ์ปกติที่ทำไว้แล้วสามารถแปลความหมายได้ตรงกับความเป็นจริงหรือไม่ เช่น นักเรียนคนหนึ่งสอบวิชาคณิตศาสตร์ได้ 20 คะแนน อันดับคะแนน T ปกติ 50 แปลว่า มีความสามารถปานกลาง ความเป็นจริงจะเป็นเช่นนั้นจริงหรือไม่ ในเรื่องนี้จึงถือว่าเป็นสิ่งสำคัญมาก

3. มีความทันสมัย เกณฑ์ปกตินั้นขึ้นอยู่กับความสามารถของประชากรกลุ่มนั้น การพัฒนาคนมีอยู่ตลอดเวลา เทคโนโลยี สภาพแวดล้อม การบริโภคอุปโภค สิ่งเหล่านี้จะช่วยให้คนเก่งหรืออ่อนได้ ดังนั้นเกณฑ์ปกติที่เคยศึกษาไว้นานแล้วหลายปี อาจจะไม่ตรงกับความเป็นจริง ๆ จึงต้องสร้างขึ้นมาใหม่ให้ทันสมัย โดยทั่วไปแล้วเกณฑ์ปกติควรเปลี่ยนทุก ๆ 5 ปี

ชนิดของเกณฑ์ปกติ แบ่งเป็น 2 ชนิด ดังนี้

1. การแบ่งชนิดของเกณฑ์ปกติตามลักษณะของประชากร ได้แก่

1.1 เกณฑ์ปกติระดับชาติ (national norms) คือ การสร้างเกณฑ์ปกติที่ใช้ประชากรทั่วประเทศ เช่น หาเกณฑ์ปกติของวิชาคณิตศาสตร์ระดับชั้นประถมศึกษาปีที่ 6 ก็ต้องสอบ นักเรียนชั้นประถมศึกษาปีที่ 6 ทั่วประเทศ จำนวนนักเรียนที่ใช้เป็นกลุ่มตัวอย่างจึงมีมากมาย

1.2 เกณฑ์ปกติระดับท้องถิ่น (local norms) เป็นการสร้างเกณฑ์ปกติระดับเล็กลงมา เช่น ระดับจังหวัด หรือระดับอำเภอ เป็นประโยชน์ในการเปรียบเทียบคะแนนของผู้สอบแต่ละคนกับคนทั้งจังหวัดหรืออำเภอ

1.3 เกณฑ์ปกติของโรงเรียน (school norms) โรงเรียนบางแห่งมีขนาดใหญ่มีนักเรียนแต่ละชั้นมีจำนวนมาก เมื่อสร้างแบบทดสอบแต่ละวิชาของแต่ละระดับชั้นจนมีคุณภาพได้มาตรฐานแล้ว จะสร้างเกณฑ์ปกติของโรงเรียนตนเองก็ได้ กรณีสร้างเกณฑ์ปกติของโรงเรียนเดียวหรือในกลุ่มโรงเรียนเดียวกัน เรียกว่า เกณฑ์ปกติของโรงเรียน ใช้ประเมินเปรียบเทียบนักเรียนแต่ละคนกับนักเรียนโดยรวมของโรงเรียน และใช้ประเมินการพัฒนาของโรงเรียนได้ด้วย โดยพิจารณาจากผลการสอบแต่ละปีว่า เติบโตหรือด้อยกว่าปีที่สร้างเกณฑ์ปกติไว้แล้ว

2. การแบ่งชนิดของเกณฑ์ปกติตามลักษณะของการใช้สถิติการเปรียบเทียบ ได้แก่

2.1 เกณฑ์ปกติเปอร์เซ็นต์ไทล์ (percentile norms) เกณฑ์ปกติแบบนี้สร้างจากคะแนนสอบที่มาจากประชากร หรือกลุ่มตัวอย่างที่เป็นตัวแทนที่ดี แล้วดำเนินการตามวิธีการสร้างเกณฑ์ปกติทั่วไป เมื่อหาค่าเปอร์เซ็นต์ไทล์เสร็จก็หยุดแค่นั้น เกณฑ์ปกติแบบนี้เป็นคะแนนจัดอันดับเท่านั้น จะนำไปบอกกลับกันไม่ได้ แต่สามารถเปรียบเทียบและแปลความหมายได้ เช่น นักเรียนคนหนึ่งสอบได้ 25 คะแนน เทียบกับเกณฑ์ปกติตรงกับตำแหน่งเปอร์เซ็นต์ไทล์ที่ 80 (PR_{80}) แสดงว่าถ้ามีคนเข้าสอบ 100 คน เขาเก่งกว่าคนอื่น 80 คน (เขาอ่อนกว่าคนอื่นเพียง 20 คน)

2.2 เกณฑ์ปกติคะแนนที (T-score norms) นิยมใช้กันมาก เพราะเป็นคะแนนมาตรฐานที่สามารถนำมาบอกกลับและหาคะแนนเฉลี่ยได้ มีความเหมาะสมในการแปลความหมายมีค่าอยู่ในช่วงคะแนน 10-90 มีคะแนนเฉลี่ย 50 ส่วนเบี่ยงเบนมาตรฐาน 10 เรียกคะแนนชนิดนี้ว่าคะแนน T ปกติ (normalized T-score)

2.3 เกณฑ์ปกติสเตโนน (stanine norms) คะแนนแบบนี้เป็นคะแนนมาตรฐานชนิดหนึ่ง แต่มีค่าเพียง 9 ตัว (standard nine point) คะแนนเฉลี่ยอยู่ที่ 5 มีส่วนเบี่ยงเบนมาตรฐานประมาณ 2 คะแนน แต่ละสเตโนนจะถูกกำหนดตามอัตราส่วนร้อยละของการแจกแจงโค้งปกติ ดังนี้

คะแนนสเตโนน	1	2	3	4	5	6	7	8	9
ร้อยละของจำนวนคนอยู่ในสเตโนน	4%	7%	12%	17%	20%	17%	12%	7%	4%

2.4 เกณฑ์ปกติตามอายุ (age norms) แบบทดสอบมาตรฐานบางอย่างหาเกณฑ์ ปกติตามอายุ เพื่อดูพัฒนาการในเรื่องเดียวกันว่า อายุต่างกันจะมีพัฒนาการอย่างไร หรืออายุเท่ากันจะมีพัฒนาการต่างกันหรือไม่ การสร้างแบบทดสอบวัดความรู้ปัญหาและความถนัด นิยมหาเกณฑ์ปกติโดยวิธีนี้ ส่วนแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนจะหาเฉพาะแบบทดสอบวิชาที่เป็นพื้นฐานจริง ๆ เช่น ภาษาและคณิตศาสตร์ เป็นต้น

2.5 เกณฑ์ปกติตามระดับชั้น (grade norms) เป็นการหาเกณฑ์ปกติตามระดับชั้นเรียนในโรงเรียน แบบทดสอบที่จะทำเกณฑ์ปกติแบบนี้ได้ต้องเป็นเนื้อหาเดียวกัน วิชาที่นิยมสร้างเกณฑ์ปกติชนิดนี้มักเป็นวิชาพื้นฐาน เช่น คำศัพท์ คณิตศาสตร์เบื้องต้น ส่วนแบบทดสอบวัดความรู้ความสามารถจะค่อนข้างกว้าง เช่น คำศัพท์ก็ให้ครอบคลุมตั้งแต่ชั้น ม.1-3 แล้วพิจารณาดูว่า ระดับชั้น ม.1, ม.2 หรือ ม.3 จะได้กี่คะแนน ก็จะเป็นเกณฑ์ปกติของระดับชั้นนั้น ๆ

ตอนที่ 3 การสังเคราะห์งานวิจัย

ในตอนที 3 นี้เป็นการนำเสนอการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องในประเด็นของการสังเคราะห์งานวิจัย โดยผู้วิจัยแบ่งหัวข้อในการนำเสนอ ซึ่งประกอบด้วย 3.1 ความหมายและความสำคัญของการสังเคราะห์งานวิจัย 3.2 ประเภทของการสังเคราะห์งานวิจัย 3.3 การวิเคราะห์อภิธาน และ 3.4 งานวิจัยที่เกี่ยวข้องกับการสังเคราะห์งานวิจัย มีรายละเอียดดังนี้

3.1 ความหมายและความสำคัญของการสังเคราะห์งานวิจัย

3.1.1 ความหมายของการสังเคราะห์งานวิจัย

การสังเคราะห์ (synthesis) หมายถึง การนำส่วนย่อยๆ หรือส่วนต่างๆ มาประกอบเข้าด้วยกันจนเกิดสิ่งใหม่ขึ้น เช่น การบรรยายความรู้สึก ประสบการณ์ โดยการนำคำต่าง ๆ มาประกอบเข้าด้วยกัน หรือการทำงานบางอย่างจนเกิดผล โดยมาจากการประชุม ระดมสมอง หรือการสร้างทฤษฎีใหม่ โดยการเชื่อมโยงทฤษฎีเก่าเข้าด้วยกัน เป็นต้น (อุทุมพร จามรมาน, 2531)

การสังเคราะห์งานวิจัย (research synthesis) หรือการปริทัศน์งานวิจัย (research review) หมายถึง การวิจัยที่เป็นการศึกษารายงานวิจัยจำนวนมากที่ศึกษาปัญหาเดียวกัน เพื่อสรุปผลการวิจัยและสรุปความคล้ายคลึงและความแตกต่างของงานวิจัยแต่ละเรื่อง รวมทั้งเป็นการอธิบายถึงสาเหตุที่มาของความแตกต่างเหล่านั้น เพื่อให้ได้องค์ความรู้ที่จะสามารถนำไปใช้ประโยชน์อย่างกว้างขวาง (Hunter, Schmidt & Jackson, 1982; Glass, McGaw, & Smith, 1981; นงลักษณ์ วิรัชชัย, 2542 อ้างถึงใน สำนักงานเลขาธิการสภาการศึกษา, 2552)

นงลักษณ์ วิรัชชัย (2542) ได้กล่าวถึงความหมายของการสังเคราะห์งานวิจัยว่า เป็นระเบียบวิธีการศึกษาหาข้อเท็จจริงเพื่อตอบปัญหาใดปัญหาหนึ่ง โดยการรวบรวมงานวิจัยเกี่ยวกับการศึกษาในเรื่องเดียวกันหลายเรื่อง มาวิเคราะห์ด้วยวิธีการทางสถิติ หรือวิธีการวิเคราะห์เนื้อหาให้ได้ข้อสรุปของผลการวิจัยเกี่ยวกับเรื่องที่ต้องสังเคราะห์

Cooper & Hedges (1994) กล่าวถึงการสังเคราะห์งานวิจัยว่า เป็นการรวบรวมงานวิจัยเชิงประจักษ์ โดยมีจุดประสงค์เพื่อการสร้างองค์ความรู้ใหม่ เป็นการพยายามค้นหาความคงเส้นคงวาและความผันแปรที่เกิดขึ้นจากการศึกษางานวิจัยที่มีลักษณะคล้ายกันโดยเป็นวิธีที่เป็นระบบ มีความกว้างครอบคลุมประเด็นที่เกี่ยวข้อง (Cooper & Hedges, 1994 อ้างถึงใน สำนักงานเลขาธิการสภาการศึกษา, 2552)

ดังนั้นอาจกล่าวได้ว่า การสังเคราะห์งานวิจัย (research synthesis) เป็นรวบรวมงานวิจัยที่ศึกษาปัญหาเดียวกันมาวิเคราะห์ สรุปด้วยวิธีการที่เป็นระบบ เพื่อให้ได้ข้อเท็จจริงหรือคำตอบที่ต้องการศึกษา ตลอดจนสร้างองค์ความรู้ใหม่หรือต่อยอดผลงานวิจัยเดิม

3.1.2 ความสำคัญของการสังเคราะห์งานวิจัย

การสังเคราะห์งานวิจัยมีความสำคัญหลายประการดังนี้ (อุทุมพร จามรมาน, 2527; สุพัฒน์ สุกมลสันต์, 2534 อ้างถึงใน ทศน์ศิริรินทร์ สว่างบุญ, 2548)

1. ในอดีตเมื่อปริมาณงานวิจัยมีไม่มากนัก นักวิจัยมักบูรณาการข้อค้นพบจากการวิจัยต่าง ๆ โดยการอ่านรายงานการวิจัยแล้วจับประเด็นที่เป็นผลสรุปของงานวิจัยออกมาในรูปการบรรยาย ผลการบูรณาการในลักษณะการบรรยายนี้มักมีความแตกต่างกันตามประเภท ตามความรู้และความเชี่ยวชาญในการจับประเด็นตลอดจนจุดเน้นของผู้สังเคราะห์

ปัจจุบันปริมาณงานวิจัยมีมากขึ้น การที่ผู้วิจัยอ่านงานวิจัยแล้วสรุปประเด็นอย่างที่เคยปฏิบัติมานั้น อาจทำให้เกิดความคลาดเคลื่อนเนื่องมาจากตัวผู้วิจัยที่ต้องอ่านงานวิจัยจำนวนมาก และไม่มีหลักเกณฑ์ที่ใช้สรุปข้อค้นพบจากงานวิจัยเหล่านั้น ดังนั้นการสังเคราะห์งานวิจัยที่เป็นระบบจึงมีความจำเป็นมากขึ้น

2. การมีงานวิจัยจำนวนมากแต่ไม่ได้นำผลการวิจัยมาใช้ ถือได้ว่าเป็นการสูญเสียเปล่าทางการลงทุนด้านเวลาและสติปัญญา ซึ่งอาจเกิดจากความไม่สมบูรณ์ของงานวิจัย สถานภาพของผู้ทำวิจัย และความจำกัดทางผลการวิจัยรวมทั้งปัจจัยอื่น ๆ

3. งานวิจัยที่ได้ทำในอดีตจนถึงปัจจุบันนี้มีความหลากหลายในเรื่องของวิธีการวิจัย เนื้อหา การวิจัย กลุ่มตัวอย่างหรือการวิเคราะห์ข้อมูล ทำให้ไม่เกิดผลกระทบ (impact) ต่อการเปลี่ยนแปลงได้อย่างจริงจัง และอาจมีงานวิจัยจำนวนหนึ่งที่มีเนื้อหาการวิจัยที่เหมือนกันหรือคล้ายคลึงกัน แต่ยังไม่มีการรวบรวมและสรุปว่าผลการวิจัยที่แท้จริงเป็นอย่างไร ทำให้เกิดความไม่มั่นใจในการนำผลการวิจัยไปประยุกต์ใช้

4. จำนวนงานวิจัยที่เพิ่มขึ้นมากมายและรวดเร็ว อาจเกิดความซ้ำซ้อนในการทำวิจัยได้

5. งานวิจัยจำนวนมากที่ศึกษาปัญหาเดียวกันโดยใช้รูปแบบการวิจัยคล้ายกันแต่ใช้กลุ่มตัวอย่างต่างกัน ผลการวิจัยมีทั้งสอดคล้องกันและขัดแย้งกัน ทำให้ผู้อื่นที่สนใจหรือผู้ที่ต้องการใช้ประโยชน์จากผลงานวิจัยเกิดความสับสนและไม่สามารถหาข้อสรุปได้

ดังนั้นจะเห็นได้ว่า การสังเคราะห์งานวิจัยมีความสำคัญและจำเป็นอย่างยิ่งในการหาข้อสรุปหรือข้อเท็จจริงที่เป็นประโยชน์ ตลอดจนการนำผลการวิจัยไปใช้เพื่อให้เกิดประโยชน์ ซึ่งจะเห็นว่ามีการพยายามพัฒนาวิธีการวิเคราะห์งานวิจัยให้เป็นระบบ มีความเที่ยงตรงและเชื่อถือได้มาอย่างต่อเนื่อง

3.2 ประเภทของการสังเคราะห์งานวิจัย

การสังเคราะห์การวิจัยแบ่งเป็น 2 แบบ แบบแรกเป็นการสังเคราะห์งานวิจัยในฐานะเป็นส่วนหนึ่งของการวิจัย การสังเคราะห์งานวิจัยแบบนี้เป็นการสังเคราะห์งานวิจัยในขั้นตอนของการศึกษารายงานวิจัยที่เกี่ยวข้องกับการวิจัย แบบที่สองเป็นการสังเคราะห์งานวิจัยที่เป็นงานวิจัยเสร็จสิ้นสมบูรณ์ในตัว ผลจากการสังเคราะห์งานวิจัยแบบนี้เป็นประโยชน์ต่อการหาคำตอบปัญหาวิจัยที่เป็นข้อสรุปที่มีความลึกซึ้งนำไปใช้ให้เกิดประโยชน์ต่อวงการวิชาการ และสังคมได้อย่างกว้างขวาง (Light & Pillermer, 1984; Glass, McGaw & Smith, 1981; Mullen, 1989; นงลักษณ์ วิรัชชัย, 2543; สำนักงานเลขาธิการสภาการศึกษา, 2552)

ในการสังเคราะห์งานวิจัยมีวิธีการดำเนินทั้งเชิงปริมาณและเชิงคุณภาพ โดยวิธีการที่นิยมนำมาใช้ในการวิจัยมี 4 วิธี มีรายละเอียดดังนี้ (นงลักษณ์ วิรัชชัย, 2542; Koricheva, et al., 2013 อ้างถึงใน สรรัญญา จันทรชูสกุล, 2561)

1) การสังเคราะห์งานวิจัยแบบพรรณนา (narrative reviews) เป็นวิธีการสังเคราะห์งานวิจัยเชิงคุณภาพโดยการเปรียบเทียบผลการวิจัยแต่ละเรื่องว่ามีส่วนที่เหมือนหรือต่างกันอย่างไร จากนั้นจึงพิจารณาว่าความแตกต่างที่เกิดขึ้นนั้นเป็นเพราะงานวิจัยมีความแตกต่างกันลักษณะใด มีการกำหนดคุณลักษณะเพื่อใช้ในการสังเคราะห์ อย่างไรก็ตามวิธีการนี้เป็นวิธีการที่ไม่มีระบบ ประสบการณ์ของนักวิจัยมีผลต่อผลการสังเคราะห์ เนื่องจากมีความเป็นอัตนัยสูง ให้ผลการสังเคราะห์ซ้ำได้คงเดิมค่อนข้างต่ำ

2) วิธีการนับคะแนน (vote counting) เป็นวิธีการสังเคราะห์งานวิจัยเชิงปริมาณ โดยการจัดกลุ่มผลการวิจัยออกเป็น 3 กลุ่มคือ กลุ่มที่มีนัยสำคัญทางสถิติและค่าสถิติเป็นบวก กลุ่มที่มีนัยสำคัญทางสถิติ

และค่าสถิติเป็นลบ และกลุ่มที่ไม่พบนัยสำคัญทางสถิติ แล้วจึงทำการสังเคราะห์โดยการนับความถี่ของผลการวิจัยในแต่ละกลุ่ม

3) วิธีรวมสะสมค่าความน่าจะเป็น (combining probabilities) เป็นวิธีการสังเคราะห์งานวิจัยเชิงปริมาณ โดยการนำค่าความน่าจะเป็น หรือระดับนัยสำคัญทางสถิติ ซึ่งเป็นดัชนีมาตรฐานที่ไม่มีหน่วย มารวมกันแล้วนำมาหาค่าเฉลี่ย ประมาณค่าความคลาดเคลื่อนมาตรฐานและทดสอบสมมติฐานว่าค่าเฉลี่ยของพารามิเตอร์ความน่าจะเป็นแตกต่างจากศูนย์หรือไม่ วิธีการนี้ไม่ยุ่งยาก และไม่มีข้อตกลงเบื้องต้นเกี่ยวกับค่าพารามิเตอร์ แต่จุดอ่อนของวิธีนี้คือ ผลการสังเคราะห์ไม่ให้ค่าประมาณค่าขนาดอิทธิพล และผลการวิเคราะห์อาจไม่ถูกต้องหากมีงานวิจัยเพียงเรื่องเดียวมีค่าพารามิเตอร์ไม่เท่ากับศูนย์

4) การวิเคราะห์อภิมาน (meta-analysis) เป็นวิธีการสังเคราะห์งานวิจัยเชิงปริมาณวิธีหนึ่ง ที่ได้รับความนิยมนำมาใช้ในการสังเคราะห์งานวิจัย เนื่องจากการวิเคราะห์อภิมานใช้สำหรับเปรียบเทียบใน ส่วนที่สามารถเปรียบเทียบได้อย่างแจ่มชัด ในรูปแบบของข้อความ จำนวน ลำดับความสำคัญ และความสัมพันธ์ระหว่างตัวแปร โดยใช้การประมาณค่าออกมาในรูปแบบของดัชนีมาตรฐาน ได้แก่ ค่าขนาดอิทธิพล (effect size: d) และค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient: r) รวมถึงเป็นการรวบรวมเอาข้อมูลที่เป็นรูปแบบที่มีอยู่หลากหลายให้เป็นรูปแบบที่ดีที่สุดจากการประมวลผล ทั้งได้ค่าประมาณที่รวบรวมจากข้อมูลหลายแหล่ง และมีการรายงานการศึกษาที่ระบุรายละเอียดที่มีการตีพิมพ์ และหลักฐานที่อ้างอิงได้ (Lam & Munzner, 2008 อ้างถึงใน นิตยภัต สุขเจริญ และ วิญญูวัฒน์ อยู่ในศิลป; 2557)

จากการสังเคราะห์งานวิจัยทั้ง 4 ประเภท สามารถเปรียบเทียบคุณลักษณะของการสังเคราะห์ได้ ดังตารางที่ 2.2 มีรายละเอียดดังนี้

ตารางที่ 2.2 การเปรียบเทียบวิธีการสังเคราะห์งานวิจัย (Koricheva, et al., 2013 อ้างถึงใน สรัญญา จันทร์ชูสกุล, 2561)

คุณลักษณะของการสังเคราะห์	การสังเคราะห์งานวิจัยแบบพรรณนา	วิธีการนับคะแนน	วิธีรวมสะสมค่าความน่าจะเป็น	การวิเคราะห์อภิมาน
1. การควบคุมประเภทของการศึกษาเป็นสิ่งที่สามารถทำได้ในการทบทวนเอกสาร	ไม่ใช่	ไม่ใช่	ไม่ใช่	ใช่
2. การแปลผลการการศึกษาโดยใช้นัยสำคัญของสถิติ	ใช่	ใช่	ใช่	ไม่ใช่
3. ขนาดของตัวอย่างและอำนาจทดสอบทางสถิติของแต่ละการศึกษาได้นำมารวมกัน	ไม่ใช่	ไม่ใช่	ใช่	ใช่
4. มีการประเมินนัยสำคัญทางสถิติของค่าเฉลี่ยอิทธิพล	ไม่ใช่	ไม่ใช่	ใช่	ใช่
5. มีการประเมินขนาดของค่าเฉลี่ยอิทธิพล	ไม่ใช่	ไม่ใช่	ไม่ใช่	ใช่
6. มีการวิเคราะห์แหล่งความแปรปรวน	ไม่ใช่	ไม่ใช่	ไม่ใช่	ใช่

อย่างไรก็ดี การวิเคราะห์ห่อภิมาณ (meta-analysis) มีลักษณะที่แตกต่างจากการสังเคราะห์งานวิจัยประเภทอื่น สามารถสรุปได้ 5 ประการ ดังนี้ (นงลักษณ์ วิรัชชัย, 2542)

1) ข้อมูลที่นำมาใช้สำหรับการสังเคราะห์ห่อภิมาณ คือ ผลการวิจัยและคุณลักษณะงานวิจัย ด้วยเหตุนี้ การนิยามตัวแปรคุณลักษณะที่ใช้ในการวิจัยต้องมีความครอบคลุมงานวิจัยแต่ละเรื่อง นอกจากนี้ ตัวแปรคุณลักษณะงานวิจัยที่ใช้ในการวิจัยแต่ละเรื่องเหมือนหรือแตกต่างกันอย่างไร ในการวิเคราะห์ห่อภิมาณต้องมีการกำหนดนิยามตัวแปรสำหรับการสังเคราะห์งานวิจัย และกำหนดตัวแปรคุณลักษณะงานวิจัย เพื่อนำไปใช้ในการวิเคราะห์ข้อมูลทางสถิติต่อไป

2) เครื่องมือที่ใช้ในการวิเคราะห์ห่อภิมาณคือ แบบบันทึกคุณลักษณะงานวิจัยและคู่มือลงรหัส ที่ต้องผ่านการตรวจสอบคุณภาพ และในกรณีที่มีผู้ลงรหัสข้อมูลตั้งแต่ 2 คนขึ้นไป จำเป็นต้องมีการตรวจสอบความสอดคล้องของผู้ลงรหัสข้อมูลด้วย

3) ในการวิเคราะห์ข้อมูลจะต้องนำผลการวิจัยที่ได้จากแต่ละเรื่องมาประมาณค่าสถิติให้อยู่ในรูปของหน่วยมาตรฐาน หรือที่เรียกว่า ดัชนีมาตรฐาน ซึ่งในการวิเคราะห์ห่อภิมาณนิยมใช้ดัชนีมาตรฐานที่เรียกว่า ขนาดอิทธิพล (effect size) เป็นค่าสถิติที่บอกถึงปริมาณผลของตัวแปรจัดกระทำ (treatment) ที่มีต่อตัวแปรตามในการวิจัยเชิงทดลอง และสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) เป็นค่าสถิติที่บอกขนาดของความแปรผันร่วมกันระหว่างตัวแปรสองตัวในการวิจัยเชิงสหสัมพันธ์ ค่าดัชนีทั้งสองสามารถแปลงค่ากลับไปกลับมาได้ จึงสามารถนำงานวิจัยเชิงทดลองและการวิจัยเชิงสหสัมพันธ์มาใช้ในการสังเคราะห์ร่วมกันได้

4) แบบแผนการวิจัย (research design) ของการวิเคราะห์ห่อภิมาณเทียบเคียงได้กับการวิจัยเชิงสำรวจ โดยมีวัตถุประสงค์เพื่ออธิบายความแตกต่างของปรากฏการณ์ ซึ่งในการวิเคราะห์ห่อภิมาณเป็นการอธิบายว่าความแตกต่างของดัชนีมาตรฐานว่าเกิดขึ้นจากตัวแปรคุณลักษณะใด

5) ตัวแปรตามในการวิเคราะห์ห่อภิมาณ คือ ผลการวิจัยที่เป็นความสัมพันธ์ระหว่างตัวแปรจากงานวิจัยที่นำมาสังเคราะห์ในรูปแบบของดัชนีมาตรฐาน ส่วนตัวแปรต้นคือ ตัวแปรคุณลักษณะงานวิจัย ซึ่งหัวใจของการวิเคราะห์ห่อภิมาณเป็นการเปรียบเทียบว่า ความสัมพันธ์ระหว่างตัวแปรที่สนใจศึกษาในแต่ละเรื่องมีค่าเป็นเท่าไรในเงื่อนไขที่แตกต่างกัน ลักษณะความสัมพันธ์มีเงื่อนไขที่ได้จากการวิเคราะห์ห่อภิมาณ ทำให้ผลการวิเคราะห์ห่อภิมาณได้ผลที่ลึกซึ้งกว่าการวิจัยเดิม

3.3 การวิเคราะห์ห่อภิมาณ

การวิเคราะห์ห่อภิมาณมาจากรากศัพท์ภาษาอังกฤษว่า meta-analysis ซึ่งบัญญัติโดย Glass ในปี ค.ศ. 1976 ในความหมายของการวิเคราะห์ผลการวิเคราะห์ (analysis of analysis) ซึ่งหมายถึง การวิเคราะห์งานวิจัยเชิงปริมาณรูปแบบหนึ่งที่มีการรวบรวมงานวิจัยหลายเรื่องที่ศึกษาปัญหาเดียวกัน มาศึกษาวิเคราะห์ด้วยวิธีการทางสถิติ เพื่อให้ได้ข้อสรุปของผลการวิจัยที่ถูกต้อง เทียบตรง และน่าเชื่อถือ โดยมุ่งศึกษาข้อมูล 2 ส่วน คือ ผลการวิจัยที่วัดออกมาในรูปดัชนีมาตรฐาน (standard indices) ซึ่งได้แก่ ขนาดอิทธิพล (effect size) ในงานวิจัยเชิงทดลอง และค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) ในงานวิจัยเชิงสหสัมพันธ์ และข้อมูลอีกส่วนหนึ่งคือข้อมูลคุณลักษณะงานวิจัยที่นำมาสังเคราะห์ (Glass, 1976 อ้างถึงใน สำนักงานเลขาธิการสภาการศึกษา, 2552)

3.3.1 ประเด็นที่ควรพิจารณาในการวิเคราะห์ห่อถัก

จากคุณลักษณะของการวิเคราะห์ห่อถักดังกล่าวข้างต้น มีประเด็นที่ควรพิจารณาในการวิเคราะห์ห่อถัก 3 ประเด็น ได้แก่ 1) งานวิจัยที่นำมาสังเคราะห์ต้องเป็นงานวิจัยที่ศึกษาปัญหาเดียวกัน 2) ข้อมูลสำหรับการวิเคราะห์ห่อถัก และ 3) หน่วยการวิเคราะห์ รายละเอียดดังนี้ (นงลักษณ์ วิรัชชัย, 2542)

1) งานวิจัยที่นำมาสังเคราะห์ต้องเป็นงานวิจัยที่ศึกษาปัญหาเดียวกัน

งานวิจัยที่นำมาสังเคราะห์ต้องมีปัญหาการวิจัยที่ตรงกัน แต่งานวิจัยนั้นอาจใช้แผนแบบการวิจัยต่างกัน เครื่องมือวัดต่างกัน กลุ่มตัวอย่างต่างกัน วิธีการวิเคราะห์ต่างกันก็ได้

2) ข้อมูลสำหรับการวิเคราะห์ห่อถัก

ข้อมูลสำหรับการวิเคราะห์ห่อถักประกอบด้วย 1) ผลการวิจัยจากงานวิจัยแต่ละเรื่อง และ 2) คุณลักษณะงานวิจัย ดังนี้

2.1) ผลการวิจัยจากงานวิจัย เนื่องจากงานวิจัยแต่ละเรื่องศึกษาตัวแปรต่างกัน หรือแม้จะศึกษาตัวแปรเหมือนกันแต่ก็อาจจะใช้เครื่องมือวัดต่างกัน ดังนั้นผลการวิจัยจากงานวิจัยแต่ละเรื่องจึงไม่อยู่ในสเกลเดียวกัน ทำให้ไม่สามารถเปรียบเทียบกันได้ จำเป็นต้องทำให้อยู่ในรูปดัชนีมาตรฐาน (standard indices) ซึ่งแยกเป็น 2 ประเภท ตามประเภทของงานวิจัย คือ ค่าขนาดอิทธิพล (effect size: d) ซึ่งเป็นดัชนีมาตรฐานสำหรับงานวิจัยเชิงทดลอง ซึ่งเป็นค่าที่ใช้เปรียบเทียบอิทธิพลของตัวแปรต้นที่มีต่อตัวแปรตามระหว่างกลุ่มทดลองกับกลุ่มควบคุม และค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient: r) ซึ่งเป็นดัชนีสำหรับงานวิจัยเชิงสหสัมพันธ์

2.2) คุณลักษณะงานวิจัย แบ่งเป็น 3 กลุ่ม ดังนี้

กลุ่มที่ 1 ตัวแปรเกี่ยวกับลักษณะการพิมพ์ ได้แก่ ปีที่พิมพ์ จำนวนหน้า คุณภาพการพิมพ์ หน่วยงานต้นสังกัดของผู้วิจัย เป็นต้น

กลุ่มที่ 2 ตัวแปรเกี่ยวกับเนื้อหาสาระ ได้แก่ ประเภททฤษฎีที่ใช้ การตั้งสมมติฐานวิจัย ลักษณะกรอบความคิดในการวิจัย ความลึกซึ้งของปัญหาการวิจัย จำนวนเอกสารอ้างอิง จำนวนเอกสารอ้างอิงเฉพาะส่วนที่เป็นงานวิจัยที่ทันสมัย ประเภทของตัวแปรต้น ประเภทของตัวแปรตาม ประเภทของกลุ่มตัวอย่าง จำนวนตัวแปรต้น จำนวนตัวแปรตาม เป็นต้น

กลุ่มที่ 3 ตัวแปรเกี่ยวกับวิธีวิทยาการวิจัย ได้แก่ ขนาดกลุ่มตัวอย่าง ค่าความเชื่อมั่นของเครื่องมือวัด ลักษณะแผนแบบการวิจัย วิธีการเลือกกลุ่มตัวอย่าง วิธีการควบคุมตัวแปรแทรกซ้อน ระยะเวลาการทดลอง ลักษณะผู้ทำการทดลอง ประเภทสถิติวิเคราะห์ที่ใช้ การตรวจสอบข้อตกลงเบื้องต้น เป็นต้น

3) หน่วยการวิเคราะห์

หน่วยการวิเคราะห์ (unit of analysis) สำหรับการวิเคราะห์ห่อถัก มีลักษณะแตกต่างจากหน่วยการวิเคราะห์ในงานวิจัยทั่วไป คือ งานวิจัยทั่วไปทางสังคมศาสตร์และพฤติกรรมศาสตร์ หน่วยการวิเคราะห์ คือ หน่วยตัวอย่างแต่ละคนที่ให้ข้อมูลสำหรับการวิจัย แต่สำหรับการวิเคราะห์ห่อถักนั้นหน่วยการวิเคราะห์จะเป็นงานวิจัยหรือการทดสอบสมมติฐาน เช่น ในการวิจัยเชิงทดลองที่ต้องการเปรียบเทียบผลสัมฤทธิ์ทางการเรียน ความคงทนในการเรียนรู้ และเจตคติต่อการเรียนระหว่างกลุ่มที่เรียนแบบปฏิบัติการกับการเรียนแบบปกติ ซึ่งมีตัวแปรตาม 3 ตัว และในการเสนอรายงานการวิจัย จะต้องวิเคราะห์ข้อมูลเปรียบเทียบความแตกต่างของตัวแปรทั้ง 3 ตัว ระหว่างกลุ่มทดลองและกลุ่มควบคุม ผลจากการวิเคราะห์ดังกล่าวจะได้ค่าขนาดอิทธิพล 3 ค่า สำหรับตัวแปรตามแต่ละตัว กล่าวได้ว่างานวิจัยเรื่องนี้มีค่า

ขนาดอิทธิพล 3 ค่า หน่วยการวัดของขนาดอิทธิพลจึงเป็นการทดสอบสมมติฐานแต่ละชุด (นงลักษณ์ วิรัชชัย, 2542)

3.2 วิธีการวิเคราะห์อภิมาน

วิธีการวิเคราะห์อภิมานที่มีการนำมาใช้ในปัจจุบันมีอยู่ 6 แนวคิดหลัก ซึ่งแต่ละวิธีลักษณะรายละเอียดของกระบวนการในการสังเคราะห์ที่แตกต่างกัน ซึ่งใช้เรียกชื่อตามผู้พัฒนา ได้แก่ วิธีของ Glass, วิธีของ Hunter, วิธีของ Slavin, วิธีของ Mullen วิธีของ Hedges & Olkin และวิธีของ Rosenthal แต่ละวิธีมีลักษณะที่สำคัญสรุปได้ดังนี้ (นงลักษณ์ วิรัชชัย, 2542; สำนักเลขาธิการสภาการศึกษา, 2552)

3.2.1 วิธีการวิเคราะห์อภิมานตามแนวคิดของ Glass เป็นวิธีที่สามารถนำมาใช้สังเคราะห์งานวิจัยเชิงทดลองและงานวิจัยเชิงสหสัมพันธ์ได้ เนื่องจากจุดเด่นของวิธีการของ Glass คือ มีสูตรในการประมาณค่าดัชนีมาตรฐาน ทั้งค่าขนาดอิทธิพลที่ได้จากงานวิจัยที่มีแบบแผนในการทดลองทุกรูปแบบ รวมถึงมีสูตรการประมาณค่าสัมประสิทธิ์สหสัมพันธ์ของ Pearson ที่ได้จากค่าสัมประสิทธิ์สหสัมพันธ์แบบอื่น ๆ นอกจากนี้ วิธีของ Glass สนใจที่จะอธิบายความแปรปรวนของค่าดัชนีมาตรฐานด้วยการวิเคราะห์ความแปรปรวน และการวิเคราะห์ถดถอย โดยมีดัชนีมาตรฐานเป็นตัวแปรตาม และตัวแปรคุณลักษณะงานวิจัยหรือตัวแปรปรับเป็นตัวแปรต้น

3.2.1 วิธีการวิเคราะห์อภิมานตามแนวคิดของ Hunter เป็นวิธีที่ให้ความสำคัญกับงานวิจัยเชิงสหสัมพันธ์มากกว่างานวิจัยเชิงทดลอง โดยมีหลักในการปรับแก้ลดความแปรปรวนจากความคลาดเคลื่อนของดัชนีมาตรฐาน เพื่อให้เหลือแต่ความแปรปรวนที่มีระบบ ก่อนที่จะทำการวิเคราะห์ต่อไปว่าความแปรปรวนที่เกิดขึ้นมาจากตัวแปรคุณลักษณะงานวิจัยหรือตัวแปรปรับตัวใด ความคลาดเคลื่อนสามารถแบ่งออกได้เป็น 5 ประเภท คือ ความคลาดเคลื่อนในการวัด (measurement error), ความคลาดเคลื่อนเนื่องจากขีดจำกัดของพิสัย (error due to range restriction), ความคลาดเคลื่อนเนื่องจากการสุ่มตัวอย่าง (sampling error) ทั้งนี้ความคลาดเคลื่อนทั้ง 3 ประเภทดังกล่าวเป็นความคลาดเคลื่อนที่สามารถปรับแก้ได้ด้วยวิธีการทางสถิติตามหลักการของ Hunter ก่อนที่จะนำข้อมูลไปวิเคราะห์ ในขณะที่ความคลาดเคลื่อนจากการคำนวณ (computation error) และความคลาดเคลื่อนจากการพิมพ์ (typographical error) เป็นความคลาดเคลื่อนที่นักวิจัยต้องป้องกันไม่ให้เกิดขึ้นในการดำเนินการวิจัย

3.2.3 วิธีการวิเคราะห์อภิมานตามแนวคิดของ Slavin เป็นแนวคิดที่ให้ความสำคัญกับงานวิจัยที่นำมาใช้ในการสังเคราะห์ โดย Slavin มีความเชื่อว่าควรสังเคราะห์จากหลักฐานที่ดีที่สุด (best evidence synthesis) ด้วยเหตุนี้จึงให้ความสำคัญกับการคัดเลือกงานวิจัยที่นำมาใช้สังเคราะห์ โดยมีการประเมินคุณภาพงานวิจัย และคัดเลือกเฉพาะเล่มที่มีคุณภาพตามเกณฑ์เพื่อนำมาใช้ในการสังเคราะห์ วิธีของ Slavin จึงมีข้อจำกัดหากงานวิจัยที่นำมาสังเคราะห์มีไม่มากทำให้การสรุปอ้างอิงมีข้อจำกัด

3.2.4 วิธีการวิเคราะห์อภิมานตามแนวคิดของ Mullen เป็นวิธีที่ใช้การประมาณค่าขนาดอิทธิพลตามแนวคิดของ Cohen ได้เสนอไว้ในปี 1969 โดยใช้ส่วนเบี่ยงเบนมาตรฐานรวมเป็นตัวหาร และใช้การประมาณค่าความสัมพันธ์ระหว่างตัวแปรโดยใช้คะแนน พิซเซอร์ ซี ในการวิเคราะห์ การประมาณค่าเฉลี่ยของดัชนีมาตรฐานใช้การหาค่าเฉลี่ยแบบถ่วงน้ำหนัก มีโปรแกรมคอมพิวเตอร์เพื่อการวิเคราะห์อภิมานชื่อ BASIC meta-analysis โปรแกรมสามารถทำงานได้มีประสิทธิภาพ โดยเฉพาะในการทดสอบความเป็นเอกพันธ์ของดัชนีมาตรฐาน การตรวจสอบว่าตัวแปรปรับตัวใดสามารถอธิบายความแปรปรวนในดัชนีมาตรฐานได้ดีที่สุด

3.2.5 วิธีการวิเคราะห์ห่อภิมาณตามแนวคิดของ Hedges & Olkin เป็นวิธีที่ใช้การประมาณค่าขนาดอิทธิพลตามแนวคิดของ Cohen วิธีนี้ให้ความสำคัญกับคุณสมบัติทางสถิติของดัชนีมาตรฐาน โดยเฉพาะค่าขนาดอิทธิพล มีการปรับแก้ความคลาดเคลื่อนตามแนวของ Hunter แต่ไม่มีการปรับแก้ความคลาดเคลื่อนในการสุ่มกลุ่มตัวอย่าง ในการสังเคราะห์ Hedges & Olkin ได้เสนอสูตรการคำนวณค่าเฉลี่ย โดยการถ่วงน้ำหนักและเสนอว่าการแจกแจงของขนาดอิทธิพลเป็นการแจกแจงแบบโค-สแควร์ ก่อนการสังเคราะห์งานวิจัย

3.2.6 วิธีการวิเคราะห์ห่อภิมาณตามแนวคิดของ Rosenthal เป็นวิธีที่นำค่าระดับนัยสำคัญมาใช้ในการวิเคราะห์และคำนวณค่าอิทธิพลได้ 2 แบบ คือ แบบแรกคำนวณจากกลุ่มตัวอย่างและค่าสถิติในการทดสอบสมมติฐาน และแบบที่สองคำนวณจากกลุ่มตัวอย่างและระดับนัยสำคัญ โดยการประมาณค่าขนาดอิทธิพลของ Rosenthal ใช้แนวคิดของ Cohen ที่ใช้ส่วนเบี่ยงเบนมาตรฐานรวมเป็นตัวหาร

การวิเคราะห์ห่อภิมาณงานวิจัยด้านนวัตกรรมการพัฒนาความคิดสร้างสรรค์ของนักเรียนระดับประถมศึกษา ผู้วิจัยใช้วิธีการวิเคราะห์ห่อภิมาณตามแนวคิดของ Glass เนื่องจากสำรวจเพื่อรวบรวมงานวิจัยเพื่อนำมาสังเคราะห์นั้น พบว่า เป็นงานวิจัยประเภททดลองทั้งหมด ซึ่งจุดเด่นของวิธีการของ Glass คือ มีสูตรในการประมาณค่าดัชนีมาตรฐาน ทั้งค่าขนาดอิทธิพลที่ได้จากงานวิจัยที่มีแบบแผนในการทดลองทุกรูปแบบ จึงมีความสอดคล้องกับรูปแบบการวิจัยเหมาะสำหรับนำมาใช้ในการวิเคราะห์ห่อภิมาณในครั้งนี้

3.3 ขั้นตอนการวิเคราะห์ห่อภิมาณ

ในการวิเคราะห์ห่อภิมาณนั้นได้มีนักวิจัยได้เสนอลำดับขั้นตอนในการวิจัย มีนักวิจัยได้เสนอไว้ดังนี้ Cooper (2017) ได้เสนอขั้นตอนของการวิเคราะห์ห่อภิมาณ แบ่งเป็น 7 ขั้นตอน ดังนี้

- ขั้นที่ 1 การกำหนดปัญหา
- ขั้นที่ 2 การสืบค้นเอกสารและงานวิจัยที่เกี่ยวข้อง
- ขั้นที่ 3 การรวบรวมข้อมูลที่ได้จากการศึกษา
- ขั้นที่ 4 การประเมินคุณภาพงานวิจัย
- ขั้นที่ 5 การวิเคราะห์และการบูรณาการของผลการการศึกษา
- ขั้นที่ 6 การตีความหลักฐานข้อมูล
- ขั้นที่ 7 การนำเสนอผลการสังเคราะห์

นงลักษณ์ วิรัชชัย (2542) ได้เสนอขั้นตอนของการวิเคราะห์ห่อภิมาณ แบ่งเป็น 5 ขั้นตอน ดังนี้

- ขั้นที่ 1 การกำหนดปัญหาวิจัยและวัตถุประสงค์การวิจัย
- ขั้นที่ 2 การศึกษาเอกสารเอกสารและงานวิจัยที่เกี่ยวข้อง
- ขั้นที่ 3 การรวบรวมข้อมูล ได้แก่ การเสาะค้นงานวิจัย การกำหนดกลุ่มตัวอย่างงานวิจัยที่จะนำมาสังเคราะห์ การสร้างเครื่องมือวิจัย และการบันทึกข้อมูล
- ขั้นที่ 4 การวิเคราะห์ข้อมูล ได้แก่ การเตรียมข้อมูลสำหรับการวิเคราะห์ การตรวจสอบลักษณะการแจกแจงของตัวแปร การวิเคราะห์ข้อมูลเบื้องต้น และการวิเคราะห์เพื่อตอบคำถามวิจัย
- ขั้นที่ 5 การสรุป และนำเสนอรายงาน

นิติบดี สุขเจริญ และ วิทยุฑฒ์ อยู่ในศีล (2557) ได้เสนอขั้นตอนของการวิเคราะห์ห่อภิมาณ แบ่งเป็น 10 ขั้นตอน ดังนี้

- ขั้นที่ 1 การสร้างสมมติฐาน และระบุประเด็นการศึกษา และศึกษาเอกสารที่เกี่ยวข้อง
- ขั้นที่ 2 การสร้างขอบเขตในการศึกษา
- ขั้นที่ 3 การสร้างเครื่องมืองานวิจัย
- ขั้นที่ 4 คัดเลือกงานวิจัย สำหรับการวิเคราะห์ลักษณะงานวิจัยที่มีคุณภาพควรจะระลึกถึงเสมอ เมื่อทำการประเมินเพื่อลด bias ต่างๆ
- ขั้นที่ 5 ประเมินคุณภาพงานวิจัย พร้อมใส่รหัสสำหรับลักษณะงานวิจัย ค่าสำหรับคำนวณ effect size มีลำดับการลงรหัสควรปฏิบัติตาม
- ขั้นที่ 6 รวบรวมข้อมูล
- ขั้นที่ 7 การวิเคราะห์ข้อมูล โดยวิเคราะห์ข้อมูลพื้นฐานและเลือก moderator ที่มีประสิทธิภาพ และหาค่าความสัมพันธ์
- ขั้นที่ 8 การวิเคราะห์ห่อภิมาณ คำนวณค่าขนาดอิทธิพล (effect size) และอภิปราย
- ขั้นที่ 9 การสรุปองค์ความรู้และข้อเสนอแนะโดยการเขียนรายงานการวิเคราะห์ห่อภิมาณ
- ขั้นที่ 10 (พิเศษ) ตรวจสอบความสมบูรณ์ของการวิเคราะห์ห่อภิมาณ

จากการสังเคราะห์ขั้นตอนการวิเคราะห์ห่อภิมาณตามที่มีนักวิจัยได้เสนอลำดับขั้นตอนในการวิเคราะห์ห่อภิมาณ สามารถสังเคราะห์ขั้นตอนได้ออกเป็น 6 ขั้นตอน ได้แก่ ขั้นที่ 1 การกำหนดปัญหาวิจัย และวัตถุประสงค์การวิจัย ขั้นที่ 2 การสืบค้นเอกสารและงานวิจัยที่เกี่ยวข้อง ขั้นที่ 3 การรวบรวมข้อมูล ขั้นที่ 4 การประเมินคุณภาพงานวิจัย ขั้นที่ 5 การวิเคราะห์ข้อมูล และขั้นที่ 6 การสรุป และนำเสนอรายงาน โดยแต่ละขั้นมีรายละเอียดดังนี้ (นงลักษณ์ วิรัชชัย, 2542; นิติบดี สุขเจริญ และ วิทยุฑฒม์ อยู่ในศิล, 2557; สรรยญา จันทร์ชูสกุล, 2561; Cooper, 2017)

ขั้นที่ 1 การกำหนดปัญหาวิจัยและวัตถุประสงค์การวิจัย

ในการกำหนดปัญหาวิจัยต้องกำหนดความสัมพันธ์ระหว่างตัวแปรที่สนใจต้องการศึกษา การเลือกหัวข้อต้องเป็นประเด็นที่มีทฤษฎีและงานวิจัยรองรับ จำนวนงานวิจัยที่จะนำมาศึกษาต้องเพียงพอในการนำมาวิเคราะห์

ในขณะที่วัตถุประสงค์สำคัญสำหรับการวิเคราะห์ห่อภิมาณ คือ เพื่อสังเคราะห์ผลการวิจัยที่ยังมีข้อขัดแย้ง หรือมีจำนวนมากให้ได้คำตอบปัญหาวิจัยที่เป็นข้อสรุป เพื่อให้ได้องค์ความรู้ใหม่ทางวิชาการที่จะนำไปใช้ประโยชน์ทางวิชาการได้ และเพื่อให้ได้ข้อค้นพบที่เป็นความสัมพันธ์แบบมีเงื่อนไขที่จะสามารถนำไปใช้ประโยชน์ในการกำหนดนโยบายและการปฏิบัติจริงต่อไป

ขั้นที่ 2 การสืบค้นเอกสารและงานวิจัยที่เกี่ยวข้อง

เป็นการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องในประเด็นที่ต้องการทำการวิเคราะห์ห่อภิมาณ โดยศึกษาเกี่ยวกับทฤษฎีทั้งหมดที่เกี่ยวข้องกับการวิจัย และสังเคราะห์งานวิจัยที่เกี่ยวข้องทั้งหมด เพื่อพัฒนาเป็นกรอบแนวคิดในการวิจัย

ขั้นที่ 3 การรวบรวมข้อมูล

- 1) การสืบค้นงานวิจัย ผู้วิจัยต้องค้นหางานวิจัยอย่างพิถีพิถันจากแหล่งข้อมูลหลากหลาย แหล่งที่เป็นการศึกษาปัญหาวิจัยเดียวกัน ผู้วิจัยจำเป็นต้องกำหนดคำสำคัญ (keywords) ในการสืบค้น ซึ่งผู้วิจัยสามารถสืบค้นได้จากห้องสมุดหรือจากแหล่งออนไลน์จากคอมพิวเตอร์
- 2) การสร้างเครื่องมือวิจัย เครื่องมือที่ใช้ในการวิเคราะห์ห่อภิมาณมีจำนวน 2 ฉบับ คือ แบบประเมินคุณภาพงานวิจัย และแบบบันทึกคุณลักษณะงานวิจัย ทั้งนี้แบบประเมินคุณภาพงานวิจัย มักนิยม

สร้างเป็นแบบมาตรฐานค่า โดยมีข้อความที่บ่งบอกถึงคุณภาพงานวิจัย ดังเช่น แบบประเมินคุณภาพงานวิจัยของสำนักเลขาธิการสภาการศึกษา กระทรวงศึกษาธิการ (2552) ซึ่งพัฒนาขึ้นโดย นางลักษณ์ วิรัชชัย และคณะ ซึ่งมีเนื้อหาสาระในการประเมิน 7 ด้าน ได้แก่ ความเป็นมาและความสำคัญของปัญหาวิจัย เอกสารและงานวิจัยที่เกี่ยวข้อง วิธีดำเนินการวิจัย ผลการวิเคราะห์ข้อมูล การสรุปอภิปรายผลและข้อเสนอแนะ การนำเสนอรายงานวิจัย และคุณภาพงานวิจัย จำนวน 30 ข้อ มีลักษณะเป็นมาตรฐานค่า 5 ระดับ โดยกำหนดเกณฑ์การให้คะแนนแต่ละระดับ (rubric) ตั้งแต่ 0-4 ในขณะที่แบบบันทึกคุณลักษณะงานวิจัย มีลักษณะเป็นแบบบันทึกข้อมูลที่มีการกำหนดรหัสสำหรับการบันทึกค่าในแต่ละรายการพร้อมคู่มือลกรหัส ที่ครอบคลุมตัวแปรคุณลักษณะที่ใช้ในการวิจัย ซึ่งผู้วิจัยควรมีการนำเครื่องมือทั้งสองประเภทมาทดลองใช้ในการประเมินและบันทึกข้อมูลก่อนดำเนินการจริง และหากในกรณีที่มีผู้ประเมินและบันทึกหลายคนควรจะมีการหาค่าความสอดคล้องระหว่างผู้รวบรวมข้อมูลด้วย

3) การกำหนดกลุ่มตัวอย่างงานวิจัยที่จะนำมาสังเคราะห์ เมื่อผู้วิจัยดำเนินการรวบรวมงานวิจัยที่จะนำมาใช้เพื่อการสังเคราะห์ได้แล้วผู้วิจัยอาจทำการสังเคราะห์งานวิจัยทุกเรื่องที่สืบค้นมาได้ หรือใช้กระบวนการในการสุ่มงานวิจัยเพื่อใช้ในการศึกษา โดยการเลือกใช้การสุ่มแบบแบ่งชั้นงานวิจัย โดยกำหนดตัวแปรลักษณะงานวิจัยเป็นตัวแปรเกณฑ์ที่ใช้ในการแบ่ง แล้วจึงทำการสุ่มงานวิจัยในสัดส่วนที่เท่า ๆ กัน หรือเลือกเฉพาะงานวิจัยที่มีการตีพิมพ์เผยแพร่ เพื่อให้มั่นใจได้ว่าเป็นงานวิจัยที่มีคุณภาพ หรือเลือกศึกษาเฉพาะงานวิจัยที่มีคุณภาพ

ขั้นที่ 4 การประเมินคุณภาพงานวิจัยและการบันทึกข้อมูล

ผู้วิจัยดำเนินการประเมินคุณภาพงานวิจัย โดยการอ่านงานวิจัยเล่มนั้นโดยละเอียด แล้วจึงประเมินคุณภาพงานวิจัยตามเกณฑ์ที่กำหนดไว้ จากนั้นจึงมีการบันทึกข้อมูลและลกรหัสสำหรับงานวิจัยตามประเด็นที่กำหนดไว้ในคู่มือลกรหัส มีการคำนวณค่าดัชนีมาตรฐาน ได้แก่ ขนาดอิทธิพล (effect size) หรือค่าสัมประสิทธิ์สหสัมพันธ์ ในการประมาณค่าดัชนีมาตรฐานจำเป็นต้องใช้ข้อมูลที่ได้จากงานวิจัยแต่ละเล่ม ซึ่งอาจเสนอค่าสถิติไม่ครบถ้วน ด้วยเหตุนี้ Glass, McGaw & Smith (1981 อ้างถึงใน นางลักษณ์ วิรัชชัย , 2542) จึงได้เสนอสูตรประมาณค่าดัชนีมาตรฐาน โดยมีสัญลักษณ์ที่ใช้ ดังนี้

\bar{Y}_E, \bar{Y}_C	คือ ค่าเฉลี่ยกลุ่มทดลอง และกลุ่มควบคุม
S	คือ ส่วนเบี่ยงเบนมาตรฐาน
G	คือ คะแนนการเปลี่ยนแปลง (change score)
g	คือ คะแนนการเปลี่ยนแปลงเศษเหลือ (residual gain score)
ay	คือ คะแนนปรับแก้ตัวร่วม (score adjusted for covariate)
SS	คือ ผลรวมกำลังสองของคะแนนเบี่ยงเบน (sum of square)
MS	คือ ค่าเฉลี่ยผลรวมกำลังสองของคะแนนเบี่ยงเบน (mean square)
df	คือ องศาอิสระ (degree of freedom)
A, B	คือ ตัวแปรต้น หรือตัวแปรจัดกระทำ
d	คือ ขนาดอิทธิพล (effect size)
r	คือ สัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson' correlation coefficient)
r_{pb}	คือ สัมประสิทธิ์สหสัมพันธ์พอยท์ไบเซเรียล (point biserial correlation coefficient)
P	คือ สัดส่วน (proportion)
n	คือ ขนาดกลุ่มตัวอย่าง

วิธีการประมาณค่าจากการคำนวณโดยตรง

1) แผนแบบการวิจัยแบบสองกลุ่มวัดครั้งหลัง (two-group posttest only design) แบ่งเป็น 2 กรณี

กรณีที่ 1 การประมาณค่าขนาดอิทธิพลกรณีมีตัวแปรตามเป็นตัวแปรต่อเนื่อง ใช้สูตรประมาณค่าขนาดอิทธิพลจากค่าเฉลี่ย ดังนี้

$$d = \frac{\bar{y}_E - \bar{y}_C}{S_Y}$$

กรณีที่ 2 การประมาณค่าขนาดอิทธิพลกรณีมีตัวแปรตามไม่ต่อเนื่อง ใช้สูตรประมาณค่าขนาดอิทธิพลจากค่าเฉลี่ย ดังนี้

$$d = \frac{P_E - P_C}{\sqrt{P_C(1 - P_C)}}$$

เมื่อทราบค่าขนาดอิทธิพล จะประมาณค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันได้ดังสูตร

$$r = d \sqrt{\frac{n}{nd^2 + 4n - 8}}; \quad n = n_E + n_C$$

2) แบบแผนการวิจัยแบบสองกลุ่มวัดก่อนและหลัง (two-group, pretest posttest only design)

การประมาณค่าขนาดอิทธิพลกรณีมีตัวแปรตามเป็นตัวแปรต่อเนื่อง และใช้คะแนนการเปลี่ยนแปลง (G) คะแนนการเปลี่ยนแปลงเศษเหลือ (g) และคะแนนปรับแก้ตัวร่วม (ay) มีสูตรประมาณค่าขนาดอิทธิพล ดังนี้

$$d = \frac{\bar{G}_E - \bar{G}_C}{S_Y}; \quad S_Y = \frac{S_G}{\sqrt{2(1 - r_{EC})}}$$

$$d = \frac{\bar{g}_E - \bar{g}_C}{S_y}; \quad S_Y = \frac{S_g}{\sqrt{1 - r_{EC}^2}}$$

$$d = \frac{\bar{ay}_E - \bar{ay}_C}{S_y}; \quad S_Y = \frac{S_{ay}}{\sqrt{1 - r_{EC}^2}}$$

3) แบบแผนการวิจัยแบบแฟคตอเรียล (factorial design)

การประมาณค่าขนาดอิทธิพลในแบบแผนการวิจัยแบบแฟคตอเรียลกรณีมีตัวแปรจัดกระทำสองตัวแปรคือ A และ B ในการประมาณค่าขนาดอิทธิพลสำหรับตัวแปรจัดกระทำ A มีสูตรดังนี้

$$d = \frac{\bar{y}_{E1} - \bar{y}_{C1}}{S_Y}; \quad S_Y = \sqrt{\frac{SS_B + SS_{AB} + SS_W}{df_B + df_{AB} + df_W}}$$

กรณีที่ 1 มีการประมาณค่าขนาดอิทธิพลสำหรับตัวแปรจัดกระทำในแต่ละระดับมีสูตร ดังนี้

$$d = \frac{\bar{y}_{E1} - \bar{y}_{C1}}{S_Y}; \quad S_Y = \sqrt{MS_W}$$

4) แบบแผนการวิจัยเชิงสหสัมพันธ์ (correlation design)

การประมาณค่าขนาดอิทธิพลในแบบแผนการวิจัยเชิงสหสัมพันธ์ เป็นการประมาณค่าจากสัมประสิทธิ์สหสัมพันธ์มีสูตร ดังนี้

$$d = \sqrt{\frac{n-2}{n} \left[\frac{2r}{\sqrt{1-r^2}} \right]}$$

วิธีการประมาณค่าทางสถิติ

ค่าสถิติที่ใช้ในการประมาณค่าขนาดอิทธิพล และค่าสัมประสิทธิ์สหสัมพันธ์ เป็นค่าจากการทดสอบสมมติฐานทางสถิติ มีดังนี้

1) การประมาณค่าจากสถิติ t ในการทดสอบความแตกต่างระหว่างค่าเฉลี่ย

ค่าสถิติ t ในการทดสอบความแตกต่างระหว่างค่าเฉลี่ยมีสองกรณี คือ กลุ่มตัวอย่างเป็นอิสระต่อกัน และกลุ่มตัวอย่างที่มีความสัมพันธ์กัน มีสูตรดังนี้

$$d = t \sqrt{\frac{1}{n_E} + \frac{1}{n_C}}$$

$$d = t \sqrt{\frac{2}{n(1-r_{EC}^2)}}$$

2) การประมาณค่าจากค่าสถิติ t ในการทดสอบนัยสำคัญของค่าสัมประสิทธิ์สหสัมพันธ์

เมื่อทราบค่าสถิติ t ในการทดสอบนัยสำคัญของค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน จะสามารถประมาณค่าสัมประสิทธิ์สหสัมพันธ์และค่าขนาดอิทธิพลจากค่าสัมประสิทธิ์สหสัมพันธ์ได้จากสูตร

$$r = \frac{t}{\sqrt{t^2 + n - 2}}$$

$$d = \sqrt{\frac{n-2}{n} \left[\frac{2r}{\sqrt{1-r^2}} \right]}$$

3) การประมาณค่าจากค่าสถิติไคสแควร์

เมื่อทราบค่าสถิติไคสแควร์ในการทดสอบความสัมพันธ์ระหว่างตัวแปรจะประมาณค่าสัมประสิทธิ์สหสัมพันธ์ได้จากสูตร

$$r = \frac{\chi^2}{\chi^2 + n}$$

4) การประมาณค่าจากค่าสถิติ F

เมื่อมีการทดสอบความแตกต่างระหว่างค่าเฉลี่ยด้วยการวิเคราะห์ความแปรปรวน และทราบค่าสถิติ F จะสามารถประมาณค่าสัมประสิทธิ์สหสัมพันธ์และค่าขนาดอิทธิพลได้จากสูตร

$$r^2 = \frac{SS_B}{SS_T} = \frac{SS_B}{SS_W + SS_B} = \sqrt{\frac{F(df_B)}{F(df_W) + F(df_B)}}$$

$$d = 2 \sqrt{\frac{F(1-r_{EC}^2)(df_W-1)}{(n_E+n_C)(df_W-2)}} \text{ เมื่อการทดลองมี 2 กลุ่ม}$$

ในกรณีที่มีการวิเคราะห์ความแปรปรวนเมื่อมีกลุ่ม k กลุ่ม West & Wiratchai (1984 อ้างถึงใน นงลักษณ์ วิรัชชัย, 2542)

$$d = 2f \sqrt{\frac{3(k-1)}{k+1}}$$

$$\text{เมื่อ } f = \frac{\eta^2}{1-\eta^2} \text{ และ } \eta^2 = \frac{F(k-1)}{F(k-1) + (n-k)}$$

5) การประมาณค่าจากสถิติ Mann-Whitney U

เมื่อมีการทดสอบความสัมพันธ์ระหว่างตัวแปรด้วยสถิติ Mann-Whitney U จะประมาณค่าสัมประสิทธิ์สหสัมพันธ์พอยท์ไปซีเรียล จากสถิติ U ก่อน แล้วจึงประมาณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน จากสูตรดังต่อไปนี้

$$r = r_{pb} \sqrt{\frac{n_1 n_2}{un}}; u = \text{ordinate of unit normal distribution}$$

$$r_{pb} = 1 - \frac{2U}{n_1 n_2}; U = \text{ค่าสถิติ Mann-Whitney U}$$

จากข้อมูลข้างต้นจะเห็นได้ว่าไม่ว่าจะมีข้อมูลประเภทใด สามารถประมาณค่าขนาดอิทธิพล และค่าสัมประสิทธิ์สหสัมพันธ์ได้ และสามารถแปลงค่าระหว่างค่าขนาดอิทธิพลและค่าสัมประสิทธิ์สหสัมพันธ์ได้เช่นเดียวกัน

ขั้นที่ 5 การวิเคราะห์ข้อมูล

ในขั้นตอนการวิเคราะห์ข้อมูลเป็นการนำข้อมูลที่ได้จากขั้นตอนที่ 4 มาดำเนินการดังนี้

- 1) การเตรียมข้อมูลสำหรับการวิเคราะห์ โดยการบันทึกในโปรแกรมคอมพิวเตอร์
- 2) การตรวจสอบลักษณะการแจกแจงของตัวแปร การบรรณาธิกรณข้อมูลเพื่อตรวจสอบข้อมูลสูญหายก่อนการวิเคราะห์

3) การวิเคราะห์ข้อมูลเบื้องต้น คือ การเสนอค่าสถิติเบื้องต้นที่สำคัญของตัวแปร เพื่อตรวจสอบความสัมพันธ์ของตัวแปร ความแตกต่าง ลักษณะการกระจายของดัชนีมาตรฐาน เป็นต้น

4) การวิเคราะห์เพื่อตอบคำถามวิจัย เป็นการวิเคราะห์ผลวิจัยซึ่งวัดในรูปดัชนีมาตรฐานว่ามีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ ถ้าไม่แตกต่างกันก็สามารถสรุปผลสังเคราะห์ทั้งหมดเพื่อตอบปัญหาวิจัยได้ แต่ถ้าแตกต่างกันอย่างมีนัยสำคัญทางสถิติต้องวิเคราะห์ต่อว่าความแตกต่างหรือความแปรปรวนนั้นเกิดจากคุณลักษณะงานวิจัยด้านใดบ้าง

ขั้นที่ 6 การสรุป และนำเสนอรายงาน

การสรุปผลการวิเคราะห์และการจัดทำรายงานในการวิเคราะห์ห่อภิมาณ การสรุปผลการวิเคราะห์นั้นจะต้องมีสรุปความรู้ และการอภิปรายที่มีความเชื่อมโยงผลการวิจัยกับทฤษฎี มีการให้ข้อเสนอแนะที่เป็นประโยชน์ทั้งเชิงนโยบาย และข้อเสนอแนะทางวิชาการ ในขณะที่รายงานการวิจัยต้องมีความสมบูรณ์ ชัดเจน ซึ่งรายงานวิจัยการวิเคราะห์ห่อภิมาณมีองค์ประกอบเหมือนรายงานวิจัยทั่วไป แต่สิ่งที่แตกต่างคือบรรณานุกรมแยกออกเป็น 2 ส่วน โดยส่วนแรกเป็นบรรณานุกรมที่ใช้ในการอ้างอิง ในขณะที่บรรณานุกรมอีกส่วนจะเป็นรายการของงานวิจัยที่นำมาสังเคราะห์

จากขั้นตอนในการดำเนินการวิเคราะห์ห่อภิมาณทั้ง 6 ขั้นตอน มีขั้นตอนในการดำเนินการวิจัยที่ไม่แตกต่างไปจากวิธีการดำเนินการวิจัยโดยทั่วไป เพียงแต่ประชากรที่ใช้ในการวิจัยเป็นงานวิจัยที่อยู่ในขอบเขตปัญหาวิจัยเดียวกันที่ได้กำหนดไว้ ดังนั้นจึงทำให้ต้องมีกระบวนการสืบค้นเพื่อรวบรวมข้อมูล มีการประเมิน

คุณภาพและบันทึกข้อมูล แทนการเก็บรวบรวมข้อมูลจากบุคคลหรือหน่วยงานต่าง ๆ และมีการนำเสนอ
 บรรณานุกรม 2 ส่วน โดยมีบรรณานุกรมของงานวิจัยที่นำมาสังเคราะห์แยกส่วนต่างหาก

3.4 งานวิจัยที่เกี่ยวข้องกับการสังเคราะห์งานวิจัย

รัชนีกุล ภิญโยภานูวัฒน์ (2536) ได้วิเคราะห์อภิमानงานวิจัยประเภทวิทยานิพนธ์เกี่ยวกับการสร้าง
 และพัฒนาแบบสอบและแบบวัด ด้วยวิธีของ Glass โดยเป็นงานวิจัยในปี พ.ศ. 2513-2533 จำนวน 92
 เล่ม เครื่องมือที่ใช้ในการวิจัย คือ แบบประเมินงานวิจัย และแบบสรุปรงานวิจัย วิเคราะห์ข้อมูลด้วย สถิติเชิง
 บรรยาย สถิติทดสอบที่ การวิเคราะห์ความแปรปรวน และการวิเคราะห์ถดถอยหพหุคูณ ผลการวิจัยพบว่า
 แบบสอบและแบบวัดมีค่าเฉลี่ยความเที่ยงเท่ากับ 0.83 ค่าเฉลี่ยความตรงเท่ากับ 0.52 และมีค่าเฉลี่ยความ
 ยากง่ายและอำนาจจำแนกปานกลาง มีพิสัยความยากง่ายแคบ แต่พิสัยอำนาจจำแนกค่อนข้างกว้าง แสดง
 ว่าแบบสอบและแบบวัดมีคุณภาพโดยเฉลี่ยค่อนข้างดี ทั้งนี้ แบบสอบและแบบวัดที่มีค่าเฉลี่ยความเที่ยงสูง
 ได้แก่ แบบสอบมาตรฐาน แบบวัดพฤติกรรมการเรียนรู้ด้านจิตตพิสัย แบบสอบและแบบวัดจัดประเภทและ
 จัดอันดับ มีการกำหนดจุดมุ่งหมายและนิยามเชิงปฏิบัติการอย่างชัดเจน ลักษณะเป็นแบบปรนัยชนิด
 เลือตอบ มีจำนวนตัวเลือก 2 และ 3 ตัวเลือก มีการสุ่มตัวอย่างแบบหลายขั้นตอน และใช้วิธีวิเคราะห์
 ความเที่ยงด้วย KR 20, แบบแบ่งครึ่งข้อสอบ และแอลฟา ความแปรปรวนของความเที่ยงของแบบสอบและ
 แบบวัดคิดเป็นร้อยละ 35 เป็นผลมาจากความแตกต่างของลักษณะของแบบสอบและแบบวัด ขั้นตอนการ
 สร้าง และคุณภาพด้านอื่นของแบบสอบและแบบวัด

ไชยยศ ไพวิทยศิริธรรม (2544) ได้ทำการสังเคราะห์งานวิจัยเกี่ยวกับคุณภาพของแบบสอบหลาย
 ตัวเลือก ประกอบด้วย ค่าความเที่ยง ค่าความตรง ค่าความยาก และค่าอำนาจจำแนก ที่มีรูปแบบของแบบ
 สอบ บริบท ของแบบสอบ วิธีการตรวจให้คะแนน และประเภทของแบบสอบต่างกัน ด้วยวิธีการวิเคราะห์
 อภิमानตามแนวคิดของ Hunter ซึ่งเป็นงานวิจัยที่พิมพ์เผยแพร่ระหว่างปี พ.ศ. 2511-2542 จำนวน 96
 เล่ม วิเคราะห์ข้อมูลโดยใช้สถิติบรรยายและการวิเคราะห์ความแปรปรวน ผลการวิจัยพบว่า ค่าเฉลี่ยใน
 ภาพรวมของค่าความเที่ยงมีค่าเฉลี่ยเท่ากับ .713 ตัวแปรปรับที่อธิบายความแตกต่างของค่าความเที่ยง
 อย่างมีนัยสำคัญและค่าเฉลี่ยของกลุ่มที่มีค่าเฉลี่ยสูงสุด ได้แก่ ตัวแปรการเรียงตัวเลือก: แบบสอบที่มีการ
 เรียงตัวเลือกแบบค่ามากไปค่าน้อย (.775) การเรียงแบบสอบ: แบบสอบที่มีการเรียงจากพฤติกรรมกร
 เรียนขั้นต่ำไปขั้นสูง (.818) วิธีการตรวจให้คะแนน: แบบสอบที่มีวิธีการตอบและตรวจให้คะแนนด้วยวิธีให้
 คะแนนรายข้อต่างกันตามความสามารถของผู้สอบ (.951) และประเภทของแบบสอบ: แบบสอบเติมคำ
 (.852) ค่าเฉลี่ยในภาพรวมของค่าความตรงเท่ากับ .528 ตัวแปรปรับที่อธิบายความแตกต่างของค่าความ
 ตรงอย่างมีนัยสำคัญและค่าเฉลี่ยของกลุ่มที่มีค่าเฉลี่ยสูงสุด ได้แก่ ตัวแปรรูปแบบตัวลวง: แบบสอบที่มี
 รูปแบบตัวลวงจากคำตอบผิดของนักเรียน (.691) การเรียงแบบสอบ: แบบสอบที่มีการเรียงจากพฤติกรรม
 การเรียนขั้นต่ำไปขั้นสูง (.623) และวิธีการตรวจให้คะแนน: แบบสอบที่มีวิธีการตอบและตรวจให้คะแนน
 ด้วยวิธีของสำราญ มีแจ้ง (.756) ค่าเฉลี่ยในภาพรวมของค่าความยากเท่ากับ .474 ตัวแปรปรับที่อธิบาย
 ความแตกต่างของค่าความยากอย่างมีนัยสำคัญและค่าเฉลี่ยของกลุ่มที่มีค่าเฉลี่ยปานกลาง ได้แก่ ตัวแปร
 การเรียงแบบสอบ: แบบสอบที่มีการเรียงโดยการสุ่ม (.484) วิธีการตรวจให้คะแนน: แบบสอบที่มีวิธีการ
 ตอบและตรวจ ให้คะแนนด้วยวิธีแก้การเดา (.497) และประเภทของแบบสอบ: แบบสอบหลายตัวเลือก (.
 538) ค่าเฉลี่ยในภาพรวมของค่าอำนาจจำแนกเท่ากับ.433 ตัวแปรปรับที่อธิบายความแตกต่างของค่า
 อำนาจจำแนกอย่างมีนัยสำคัญและค่าเฉลี่ยของกลุ่มที่มีค่าเฉลี่ยสูงสุด ได้แก่ ตัวแปรรูปแบบตัวลวง: แบบ

สอบที่มีรูปแบบตัวลวงจากคำตอบผิดของนักเรียน (.550) และวิธีการตรวจให้คะแนน:แบบสอบที่มีวิธีการตอบและตรวจให้คะแนนด้วยวิธีของ อนันต์ ศรีโสภา (.693)

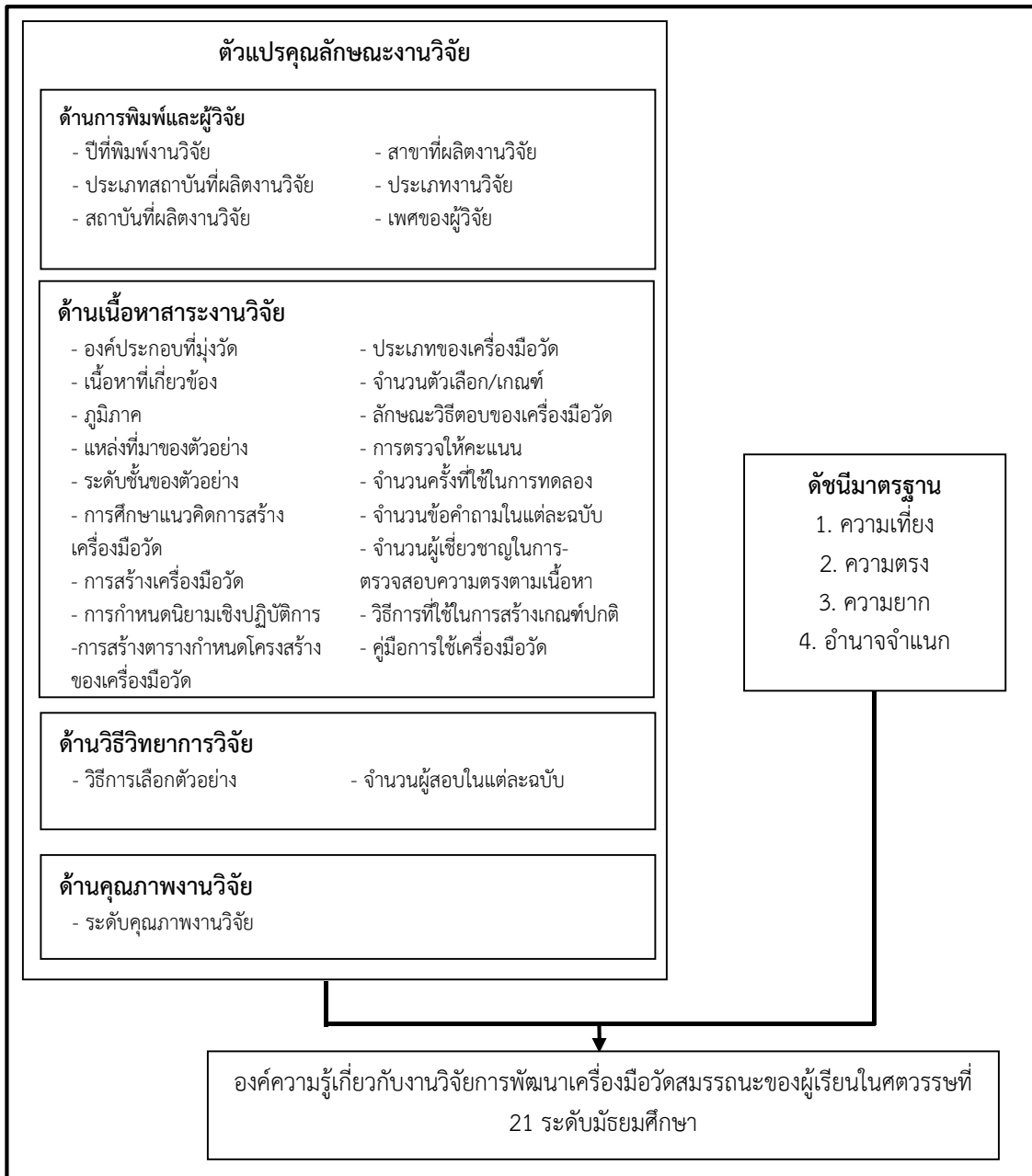
อาภาพรรณ ประทุมไทย (2553) ได้ดำเนินการวิเคราะห์ห่อภิมาณคุณภาพของแบบวัดความถนัด โดยมีวัตถุประสงค์ คือ 1) เพื่อประเมินและเปรียบเทียบคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัดระหว่างกลุ่มงานวิจัยที่มีคุณลักษณะต่างกัน และ 2) เพื่อสังเคราะห์งานวิจัยเกี่ยวกับคุณภาพของแบบวัดความถนัด ได้แก่ ค่าความตรง ค่าความเที่ยง ค่าความยาก และค่าอำนาจจำแนกที่มีคุณลักษณะงานวิจัยแตกต่างกัน ด้วยวิธีการวิเคราะห์ห่อภิมาณ งานวิจัยที่นำมาสังเคราะห์มี จำนวน 61 เล่ม ที่พิมพ์เผยแพร่ระหว่างปี พ.ศ. 2509-2551 ข้อมูลของการวิจัยครั้งนี้ประกอบด้วยค่าดัชนีมาตรฐาน จำนวน 957 ค่า การวิเคราะห์ข้อมูลใช้สถิติบรรยายและการวิเคราะห์ความแปรปรวน ผลการสังเคราะห์งานวิจัย พบว่า 1 งานวิจัยมีคุณภาพโดยรวมอยู่ในระดับปานกลาง ทั้งนี้ ตัวแปรคุณลักษณะงานวิจัยที่ทำให้เกิดความแตกต่างของค่าเฉลี่ยคุณภาพงานวิจัย จากการวิเคราะห์ความแปรปรวน โดยตัวแปรที่มีนัยสำคัญที่ระดับ .001 ได้แก่ ปีที่พิมพ์ สาขาวิชาที่ผลิตงานวิจัย ประเภทงานวิจัย การสร้างแบบวัด วิธีการเลือกกลุ่มตัวอย่าง วิธีการรวบรวมข้อมูล การศึกษาทฤษฎีวิธีการสร้างแบบวัดและการสร้างตารางกำหนดแผนผังการสร้างข้อสอบ ตัวแปรที่มีนัยสำคัญที่ระดับ .01 คือ จำนวนครั้งในการทดลองใช้ และตัวแปรที่มีนัยสำคัญที่ระดับ .05 ได้แก่ ภูมิฐานะของกลุ่มตัวอย่าง และแบบวัดความถนัดที่ใช้เป็นแนวทางในการสร้าง และ 2) ตัวแปรคุณลักษณะงานวิจัยที่ส่งผลให้ค่าความเที่ยง ค่าความตรง ค่าความยาก และค่าอำนาจจำแนกแตกต่างกัน ได้แก่ สถาบันที่ผลิตงานวิจัย กลุ่มทฤษฎีหลัก ระดับการศึกษาของกลุ่มตัวอย่าง ภูมิฐานะของกลุ่มตัวอย่าง และเวลาที่ใช้ในการสอบ

สุวิมล เสวกสุริยวงศ์ (2553) ได้การสังเคราะห์งานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ มีวัตถุประสงค์การวิจัย 3 ประการ คือ 1) เพื่อวิเคราะห์คุณลักษณะของงานวิจัยที่ศึกษาเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ 2) เพื่อประเมินคุณภาพของงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ และ (3) เพื่อสังเคราะห์องค์ความรู้ที่ได้จากงานวิจัยที่เกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ งานวิจัยที่นำมาสังเคราะห์จำนวน 75 เล่ม ที่ตีพิมพ์ในช่วงปี พ.ศ. 2523 - 2551 เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบบันทึกคุณลักษณะงานวิจัย และแบบประเมินคุณภาพงานวิจัย วิเคราะห์ข้อมูลด้วยสถิติบรรยาย ผลการวิจัยพบว่า 1) งานวิจัยส่วนใหญ่เป็นงานวิจัยระดับมหาบัณฑิต ทางภาคตะวันออกเฉียงเหนือ มีวัตถุประสงค์เพื่อสร้างแบบสอบวินิจฉัย และตรวจสอบคุณภาพเครื่องมือระดับชั้นที่มีการสร้างแบบสอบวินิจฉัยมากที่สุด คือ มัธยมศึกษาปีที่ 3 2) คุณภาพงานวิจัยส่วนใหญ่อยู่ในเกณฑ์ดีมาก 3) ผลการสังเคราะห์ความรู้ พบว่า การสร้างแบบสอบวินิจฉัยมีขั้นตอน คือ 3.1) กำหนดจุดมุ่งหมายในการสร้าง 3.2) ศึกษาทฤษฎี และเอกสารที่เกี่ยวข้องกับการสร้างแบบสอบวินิจฉัย 3.3) วิเคราะห์เนื้อหาและจุดมุ่งหมายของหลักสูตร 3.4) สร้างแบบสอบเพื่อสำรวจ 3.5) ผู้เชี่ยวชาญตรวจสอบ 3.6) ทดสอบเพื่อสำรวจจุดบกพร่องและรวบรวมคำตอบ 3.7) สร้างแบบสอบวินิจฉัย 3.8) ผู้เชี่ยวชาญตรวจสอบความตรง 3.9) ทดลองใช้แบบสอบ 3.10) วิเคราะห์ คัดเลือกและปรับปรุงข้อสอบ 3.11) วิเคราะห์คุณภาพแบบสอบและวิเคราะห์จุดบกพร่อง และ 3.12) จัดทำคู่มือการใช้แบบสอบวินิจฉัยและจัดพิมพ์เป็นรูปเล่ม

จากการศึกษางานวิจัยที่ดำเนินการสังเคราะห์งานวิจัยเกี่ยวกับการพัฒนาเครื่องมือวัด ในรูปแบบของแบบสอบและแบบวัด พบว่า สามารถดำเนินการสังเคราะห์ได้ทั้ง 2 วิธี กล่าวคือ การสังเคราะห์งานวิจัยเชิงปริมาณด้วยการวิเคราะห์ห่อภิมาณ โดยกำหนดให้ตัวแปรตามคือ ค่าดัชนีมาตรฐานคุณภาพเครื่องมือ ทั้งในส่วนของคุณค่าความตรง ค่าความเที่ยง ค่าความยาก และค่าอำนาจจำแนก แล้วใช้ตัวแปรคุณลักษณะงานวิจัยเป็นตัวอธิบายความแตกต่างที่เกิดขึ้นของคุณภาพเครื่องมือวัดดังกล่าว และอีกวิธีการหนึ่งคือ การสังเคราะห์งานวิจัยด้วยวิธีการเชิงคุณภาพ ซึ่งจะช่วยให้ได้รับองค์ความรู้เกี่ยวกับกระบวนการพัฒนาเครื่องมือวัด และข้อค้นพบในการพัฒนาเครื่องมือวัดให้มีคุณภาพต่อไป ด้วยเหตุนี้การวิจัยในครั้งนี้ ผู้วิจัยจึงได้ออกแบบการสังเคราะห์งานวิจัยครอบคลุมทั้งในมิติของการวิเคราะห์ห่อภิมาณ ซึ่งเป็นการสังเคราะห์เชิงปริมาณ และการสังเคราะห์งานวิจัยเชิงคุณภาพ ด้วยวิธีการวิเคราะห์เนื้อหา

ตอนที่ 4 กรอบแนวคิดในการวิจัย

จากการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง พบว่า ในการกำหนดตัวแปรคุณลักษณะงานวิจัยที่นิยมนำมาใช้ศึกษาประกอบด้วย ตัวแปรคุณลักษณะงานวิจัย 4 ด้าน คือ ด้านการพิมพ์และผู้วิจัย ด้านเนื้อหาสาระงานวิจัยด้านวิธีวิทยาการวิจัย และและด้านคุณภาพงานวิจัย ประกอบกับการสำรวจรายละเอียดของงานวิจัยที่นำมาใช้ในการสังเคราะห์ในเบื้องต้น ผู้วิจัยจึงได้กำหนดกรอบแนวคิดในการวิจัย ดังภาพที่ 2.1



ภาพที่ 2.2 กรอบแนวคิดในการวิจัย

บทที่ 3 วิธีดำเนินการวิจัย

การวิจัยครั้งนี้วัตถุประสงค์เพื่อ 1) ประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา 2) ศึกษาคุณลักษณะของงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา และ 3) สังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา มีวิธีดำเนินการวิจัยดังนี้

ประชากรและตัวอย่าง

ประชากร คือ งานวิจัยประเภทวิทยานิพนธ์ระดับบัณฑิตศึกษา ทั้งระดับมหาบัณฑิตและดุษฎีบัณฑิต และงานวิจัยของสถาบันต่าง ๆ โดยไม่นับรวมสารนิพนธ์หรือการค้นคว้าอิสระ ที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ประกอบด้วยสมรรถนะสำคัญได้แก่

1) ทักษะและความรู้ในเนื้อหาขั้นสูง

- 1.1) จิตสำนึกต่อโลก (global awareness)
- 1.2) ความรู้เรื่องพลเมือง (civic literacy)

2) ทักษะการคิดขั้นสูง

- 2.1) การคิดอย่างมีวิจารณญาณ (critical thinking)
- 2.2) การแก้ปัญหา (solving problems)
- 2.3) การสร้างสรรค์และนวัตกรรม (creativity and innovation)

3) ทักษะต่อตนเองและสังคม

- 3.1) การสื่อสาร (communications literacy/skill)
- 3.2) ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (collaboration, teamwork)
- 3.3) ภาวะผู้นำ (leadership)
- 3.4) ความรับผิดชอบ (responsibility)

4) ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี

- 4.1) การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy)
- 4.2) การรู้เท่าทันเทคโนโลยี (technology literacy)

โดยเป็นงานวิจัยที่มีการตีพิมพ์เผยแพร่ระหว่างปี พ.ศ. 2546-2562 เนื่องจากเป็นปีที่เริ่มมีการกำหนดกรอบแนวคิดสำหรับทักษะแห่งศตวรรษที่ 21 ภายใต้อาเซียนระหว่าง ห้องวิจัยด้านการศึกษาระดับภูมิภาคตอนเหนือ ในประเทศสหรัฐอเมริกา (North Central Regional Educational Laboratory : NCREL) และกลุ่ม Metiri ในชื่อของกรอบแนวคิด “enGauge” (Metiri Group, 2003) ผู้วิจัยดำเนินการสืบค้นและคัดเลือกงานวิจัยที่สืบค้นได้จากห้องสมุดสถาบันอุดมศึกษา และฐานข้อมูลออนไลน์ เช่น เว็บไซต์โครงการเครือข่ายห้องสมุดในประเทศไทย หรือ ThaiLIS (www.thailis.or.th/tdc/), ศูนย์ข้อมูลการวิจัย Digital "วช." (<https://dric.nrct.go.th/Index>) และเว็บไซต์ห้องสมุดของสถาบันอุดมศึกษาต่าง ๆ ผลการสืบค้นงานวิจัยพบงานวิจัยจำนวน 92 เล่ม ซึ่งเป็นประชากรในการวิจัยครั้งนี้

ตัวอย่าง คือ งานวิจัยประเภทวิทยานิพนธ์ระดับบัณฑิตศึกษา ทั้งระดับมหาบัณฑิตและดุษฎีบัณฑิต และงานวิจัยของสถาบันต่าง ๆ ที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ตามที่ได้มีการคัดลอกจากประชากร มีวิธีการได้มาซึ่งตัวอย่างในการวิจัย ดังนี้

เกณฑ์ในการคัดเลือกงานวิจัย

สำหรับการวิจัยครั้งนี้ ผู้วิจัยได้ใช้เกณฑ์ในการคัดเลือก ดังนี้

1. เป็นงานวิจัยที่สามารถสืบค้นฉบับสมบูรณ์ (full text) ได้ ทั้งในรูปแบบรูปเล่มจากห้องสมุดสถาบันอุดมศึกษา และฐานข้อมูลออนไลน์
2. เป็นงานวิจัยที่พัฒนาเครื่องมือวัดสมรรถนะตามที่คุณวิจัยกำหนด กล่าวคือ เป็นงานวิจัยที่มีวัตถุประสงค์เพื่อพัฒนาเครื่องมือวัดประเภท แบบสอบ แบบวัด และแบบประเมิน โดยไม่นับรวมงานวิจัยประเภทพัฒนารูปแบบการประเมิน รวมถึงเครื่องมือเชิงคุณภาพ และมีการรายงานค่าคุณภาพของเครื่องมือในมิติความเที่ยง ความตรง ความยาก และอำนาจจำแนก ค่าใดค่าหนึ่งเป็นอย่างน้อย

ขั้นตอนการคัดเลือกตัวอย่าง

1. สืบค้นงานวิจัยประเภทวิทยานิพนธ์ระดับบัณฑิตศึกษา ทั้งระดับมหาบัณฑิตและดุษฎีบัณฑิต และงานวิจัยของสถาบันต่าง ๆ ที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ดำเนินการสืบค้นและคัดเลือกงานวิจัยที่สืบค้นได้จากฐานข้อมูลออนไลน์ เช่น เว็บไซต์โครงการเครือข่ายห้องสมุดในประเทศไทย หรือ ThaiLIS (www.thailis.or.th/tdc/), ศูนย์ข้อมูลการวิจัย Digital "วช." (<https://dric.nrct.go.th/Index>) และเว็บไซต์ห้องสมุดของสถาบันอุดมศึกษาต่าง ๆ โดยใช้คำสืบค้นตามสมรรถนะสำคัญที่คุณวิจัยได้กำหนดไว้ และคัดเลือกเฉพาะงานวิจัยที่ตีพิมพ์เผยแพร่ระหว่างปี พ.ศ. 2546-2562 เท่านั้น โดยผู้วิจัยดำเนินการสืบค้นในเดือนสิงหาคม-ตุลาคม 2562
2. คัดเลือกงานวิจัยที่มีวัตถุประสงค์สำคัญเพื่อพัฒนาเครื่องมือวัดประเภท แบบสอบ แบบวัด และแบบประเมิน สมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา
3. สืบค้นการรายงานคุณภาพของเครื่องมือในมิติความเที่ยง ความตรง ความยาก และอำนาจจำแนก ด้วยทฤษฎีการทดสอบแบบดั้งเดิมในเบื้องต้น

ผลการสำรวจงานวิจัยในขั้นตอนที่ 1-3 พบว่า มีงานวิจัยที่เป็นไปตามเกณฑ์ข้างต้นจำนวน 90 เล่ม ซึ่งเป็นตัวอย่างในการวิจัยครั้งนี้ ผู้วิจัยดำเนินการประเมินคุณภาพงานวิจัย และบันทึกคุณลักษณะงานวิจัยเพื่อการวิเคราะห์ห่อภิมาณและการสังเคราะห์เชิงคุณภาพในลำดับต่อไป สำหรับงานวิจัยอีกจำนวน 2 เล่ม ที่ไม่ปฏิบัติตามเกณฑ์ในการวิเคราะห์ห่อภิมาณในครั้งนี้ ผู้วิจัยได้มีการประเมินคุณภาพงานวิจัยดังกล่าว และนำไปร่วมสังเคราะห์เชิงคุณภาพต่อไป

ตัวแปรที่ใช้ในการวิจัย

ตัวแปรที่ใช้ในการศึกษาครั้งนี้เป็นคุณลักษณะของงานวิจัย แบ่งออกเป็น 4 ด้าน คือ ด้านการพิมพ์และผู้วิจัย ด้านเนื้อหาสาระงานวิจัย ด้านวิธีวิทยาการวิจัย และด้านคุณภาพงานวิจัย มีรายละเอียดดังนี้

1. **ตัวแปรคุณลักษณะของงานวิจัยด้านการพิมพ์และผู้วิจัย** ประกอบด้วย 6 ตัวแปร เป็นตัวแปรจัดประเภททั้งหมด ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, ประเภทงานวิจัย และเพศของผู้วิจัย มีรายละเอียดของการกำหนดค่าตัวแปรจัดประเภท ดังนี้

1.1 ปีที่พิมพ์งานวิจัย เป็นตัวแปรการตีพิมพ์เผยแพร่่างานวิจัยระหว่างปี พ.ศ. 2546-2562 โดยระบุเลขท้าย 2 ตัวของปี พ.ศ.

1.2 ประเภทสถาบันที่ผลิตงานวิจัย เป็นประเภทของสถาบันอุดมศึกษาที่ผลิตงานวิจัย ประกอบด้วยรหัสตัวแปรดังนี้ 1 = มหาวิทยาลัยของรัฐ, 2 = มหาวิทยาลัยในกำกับของรัฐ, 3 = มหาวิทยาลัยราชภัฏ, 4 = มหาวิทยาลัยเอกชน และ 5 = สถาบันอื่น ๆ (รายงานวิจัย)

1.3 สถาบันที่ผลิตงานวิจัย เป็นรายชื่อสถาบันอุดมศึกษาที่ผลิตงานวิจัย จำนวน 16 แห่ง ประกอบด้วยรหัสตัวแปรดังนี้ 1 = จุฬาลงกรณ์มหาวิทยาลัย, 2 = มหาวิทยาลัยเกษตรศาสตร์, 3 = มหาวิทยาลัยเชียงใหม่, 4 = มหาวิทยาลัยขอนแก่น, 5 = มหาวิทยาลัยทักษิณ, 06 = มหาวิทยาลัยนเรศวร, 07 = มหาวิทยาลัยบูรพา, 08 = มหาวิทยาลัยมหาสารคาม, 09 = มหาวิทยาลัยศรีนครินทรวิโรฒ, 10 = มหาวิทยาลัยสงขลานครินทร์, 11 = มหาวิทยาลัยสุโขทัยธรรมาธิราช, 12 = มหาวิทยาลัยราชภัฏกาญจนบุรี, 13 = มหาวิทยาลัยราชภัฏนครราชสีมา, 14 = มหาวิทยาลัยราชภัฏเพชรบูรณ์, 15 = มหาวิทยาลัยราชภัฏมหาสารคาม, 16 = มหาวิทยาลัยราชภัฏเลย, 17 = มหาวิทยาลัยราชภัฏอุบลราชธานี, 18 = สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) และ 19 = สถาบันอื่น (รายงานวิจัย)

1.4 สาขาที่ผลิตงานวิจัย (MAJOR) แบ่งออกเป็น 7 สาขา ประกอบด้วยรหัสตัวแปร 1 = การวัดผลการศึกษา (การวัดและประเมินผลการศึกษา) 2 = วิจัยวิทยาการวิจัยการศึกษา, 3 = วิจัยและประเมินผลการศึกษา (การวัดผลและการวิจัยการศึกษา, วัดผลและสถิติการศึกษา), 4 = เทคโนโลยีการวัดทางการศึกษา, 5 = จิตวิทยาการแนะแนว 6 = สาขาอื่น ๆ และ 7 = แหล่งอื่น (รายงานวิจัย)

1.5 ประเภทงานวิจัย แบ่งเป็น 3 กลุ่มประกอบด้วยรหัสตัวแปรดังนี้ 1 = วิทยานิพนธ์ปริญาโท, 2 = วิทยานิพนธ์ปริญาเอก และ 3 = รายงานวิจัย

1.6 เพศของผู้วิจัย ประกอบด้วยรหัสตัวแปร ดังนี้ 1 = ชาย, 2 = หญิง และ 3 = ชายและหญิง

2. ตัวแปรคุณลักษณะของงานวิจัยด้านเนื้อหาสาระงานวิจัย ประกอบด้วย 34 ตัวแปร เป็นตัวแปรจัดประเภท 28 ตัวแปร ได้แก่ วัตถุประสงค์ของการวิจัย, เครื่องมือวัดที่พัฒนาขึ้น, แนวคิด/ทฤษฎีหลัก, องค์ประกอบที่มุ่งวัด, เนื้อหาที่เกี่ยวข้อง, ภูมิภาค, แหล่งที่มาของตัวอย่าง, ระดับชั้นของตัวอย่าง, การศึกษาแนวคิดการสร้างเครื่องมือวัด, การสร้างเครื่องมือวัด, การกำหนดนิยามเชิงปฏิบัติการ, การระบุขั้นตอนในการพัฒนาเครื่องมือวัด, การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, จำนวนครั้งที่ใช้ในการทดลอง, ทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือวัด, วัตถุประสงค์ของการทดลองใช้, ดัชนีแสดงคุณภาพ, ประเภทความเที่ยง, ประเภทความตรง, วิธีการหาค่าความยาก, วิธีการหาค่าอำนาจจำแนก, การตรวจให้คะแนน, ประเภทของเกณฑ์ปกติ, วิธีที่ใช้ในการสร้างเกณฑ์ปกติ และคู่มือการใช้เครื่องมือวัด และมีตัวแปรต่อเนื่อง จำนวน 6 ตัวแปร ได้แก่ จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา, จำนวนผู้เชี่ยวชาญด้านการวัดผล, จำนวนผู้เชี่ยวชาญด้านเนื้อหา, จำนวนข้อที่สร้างก่อนการทดลองใช้, จำนวนข้อคำถามใช้จริง และจำนวนข้อคำถามในแต่ละฉบับ ทั้งนี้มีรายละเอียดของการกำหนดค่าของตัวแปรจัดประเภท ดังนี้

2.1 วัตถุประสงค์ของการวิจัย เป็นจุดมุ่งหมายของการวิจัยที่งานวิจัยกำหนดไว้ ตั้งรหัสตัวแปรเป็นตัวแปรตมมี 0 และ 1 มีวัตถุประสงค์ของการวิจัย ดังนี้ พัฒนาตัวบ่งชี้/องค์ประกอบ, สร้าง/พัฒนาเครื่องมือวัด, ตรวจสอบคุณภาพเครื่องมือวัด, สร้างปกติวิสัย (norms), พัฒนาเกณฑ์การประเมิน/เกณฑ์การให้คะแนน, ศึกษาระดับ/เปรียบเทียบคุณลักษณะที่วัดกับตัวแปรอิสระ, ทดสอบความไม่แปรเปลี่ยนและสร้างคู่มือการใช้เครื่องมือวัด

2.2 เครื่องมือวัดที่พัฒนาขึ้น เป็นชื่อเครื่องมือวัดที่งานวิจัยพัฒนาขึ้น ประกอบด้วยรหัสตัวแปรดังนี้ 1 = เครื่องมือวัดทักษะแห่งศตวรรษที่ 21, 2 = เครื่องมือวัดความตระหนักต่อโลก, 3 = เครื่องมือวัดคุณลักษณะอันพึงประสงค์ (ความรับผิดชอบ, คุณธรรม จริยธรรม และค่านิยมที่พึงประสงค์, ความเป็น

พลเมืองดี), 4 = เครื่องมือวัดความมีวินัยในตนเอง, 5 = เครื่องมือวัดทักษะชีวิต/ทักษะชีวิตและอาชีพ, 6 = เครื่องมือวัดทักษะการคิด/ทักษะการคิดขั้นสูง, 7 = เครื่องมือวัดความสามารถการคิดอย่างมีวิจารณญาณ, 8 = เครื่องมือวัดทักษะการแก้ปัญหา (คณิตศาสตร์, วิทยาศาสตร์), 9 = เครื่องมือวัดความคิดสร้างสรรค์, 10 = เครื่องมือวัดทักษะกระบวนการทางคณิตศาสตร์, 11 = เครื่องมือวัดจิตวิทยาศาสตร์, 12 = เครื่องมือวัดความสามารถในการอ่าน คิววิเคราะห์ และเขียนสื่อความ, 13 = เครื่องมือวัดความสามารถทางด้านเทคโนโลยีสารสนเทศ (จริยธรรมในการใช้เทคโนโลยีสารสนเทศ, ความไว้วางใจระหว่างบุคคลในเครือข่ายสังคมออนไลน์, พฤติกรรมการติดอินเทอร์เน็ต) 14 = เครื่องมือวัดทักษะการรู้เท่าทันสื่อ/การรู้เท่าทันสื่อและสารสนเทศ/การรู้ดิจิทัล และ 15 = อื่น ๆ

2.3 แนวคิด/ทฤษฎีหลัก เป็นแนวคิดหรือทฤษฎีหลักสำหรับใช้พัฒนาเครื่องมือวัด ประกอบด้วยรหัสตัวแปร ดังนี้ 1 = กรอบแนวคิดการเรียนรู้ในศตวรรษที่ 21 ของ Partnership for 21st Century Learning, 2 = ความตระหนักต่อโลกตามแนวคิดของ Hanvey (1982), 3 = หลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ.2544, 4 = หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ.2551 (สมรรถนะสำคัญ/คุณลักษณะอันพึงประสงค์/ทักษะกระบวนการ), 5 = แนวคิดทักษะชีวิตของกรมสุขภาพจิต, 6 = แนวคิดการคิดอย่างมีวิจารณญาณของ Dressel & Mayhew, 7 = แนวคิดการคิดอย่างมีวิจารณญาณของ Watson & Glaser, 8 = แนวคิดการคิดอย่างมีวิจารณญาณของ Ennis (1985), 9 = ทฤษฎีความคิดสร้างสรรค์ของ Torrance, 10 = ทฤษฎีความคิดสร้างสรรค์ของ Guilford, 11 = ทฤษฎีปัญญาแห่งความสำเร็จของ Sternberg, 12 = คุณลักษณะจิตวิทยาศาสตร์ ของ สสวท., 13 = แนวคิดจริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสารของ Mason, Mason & Culnan (1995), 14 = แนวคิดจริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสารของ Mason (2001), 15 = สมรรถนะของการรู้เท่าทันสื่อและสารสนเทศ ของ UNESCO (2013), 16 = แนวคิดการรู้เท่าทันสื่อและดิจิทัล ของ Hobbs (2011), 17 = สังเคราะห์จากหลายแนวคิด/ทฤษฎี, 18 = อื่น ๆ และ 99 = ไม่ระบุ

2.4 องค์ประกอบที่มุ่งวัด เป็นองค์ประกอบของเครื่องมือวัดมุ่งวัด ตั้งรหัสตัวแปรเป็นตัวแปรเต็มมี 0 และ 1 โดยมีองค์ประกอบที่มุ่งวัด ดังนี้ จิตสำนึกต่อโลก, ความรู้เรื่องพลเมือง, การคิดอย่างมีวิจารณญาณ, การแก้ปัญหา, การสร้างสรรค์และนวัตกรรม, การสื่อสาร, ทักษะด้านการทำงานเป็นทีม ความร่วมมือ, ภาวะผู้นำ, ความรับผิดชอบ, การรู้เท่าทันสื่อและสารสนเทศ และการรู้เท่าทันเทคโนโลยี

2.5 เนื้อหาที่เกี่ยวข้อง เป็นเนื้อหาที่ของเครื่องมือวัดที่พัฒนาขึ้นเชื่อมโยงกับกลุ่มสาระการเรียนรู้ใด ประกอบด้วยรหัสตัวแปร ดังนี้ 1 = ภาษาไทย, 2 = คณิตศาสตร์, 3 = วิทยาศาสตร์, 4 = สังคมศึกษา ศาสนา และวัฒนธรรม, 5 = สุขศึกษาและพลศึกษา, 6 = ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์), 7 = การงานอาชีพและเทคโนโลยี, 8 = ภาษาต่างประเทศ และ 9 = ไม่สังกัดกลุ่มสาระ (content free)

2.6 ภูมิภาค เป็นภูมิภาคของประชากรและตัวอย่างสำหรับการวิจัย แบ่งเป็น 5 ภูมิภาค ประกอบด้วยรหัสตัวแปร ดังนี้ 1 = กรุงเทพมหานคร, 2 = ภาคเหนือ, 3 = ภาคตะวันออกเฉียงเหนือ, 4 = ภาคกลาง 5 = ภาคใต้ และ 6 = มากกว่า 1 ภูมิภาค

2.7 แหล่งที่มาของตัวอย่าง เป็นต้นสังกัดของสถานศึกษาที่ประชากรหรือตัวอย่างศึกษาอยู่ แบ่งเป็น 6 แหล่ง ประกอบด้วยรหัสตัวแปร ดังนี้ 1 = สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.), 2 = สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน, 3 = องค์กรปกครองส่วนท้องถิ่น (อปท.), 4 = สำนักงานการศึกษา กรุงเทพมหานคร (กทม.), 5 = สำนักงานคณะกรรมการการอุดมศึกษา (สกอ.) 6 = อื่นๆ และ 9 = ไม่ระบุ

2.8 ระดับชั้นของตัวอย่าง แบ่งเป็น 6 ชั้นปี ตามระดับชั้นมัธยมศึกษา ประกอบด้วยรหัสตัวแปร ดังนี้ 1 = ชั้นมัธยมศึกษาปีที่ 1, 2 = ชั้นมัธยมศึกษาปีที่ 2, 3 = ชั้นมัธยมศึกษาปีที่ 3, 4 = ชั้นมัธยมศึกษาปีที่ 4, 5 = ชั้นมัธยมศึกษาปีที่ 5, 6 = ชั้นมัธยมศึกษาปีที่ 6, 7 = ชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน), 8 = ชั้นมัธยมศึกษาตอนปลาย (มากกว่า 1 ชั้นเรียน), 9 = ชั้นมัธยมศึกษาตอนต้นและตอนปลาย และ 10 = ชั้นประถมศึกษาและมัธยมศึกษา

2.9 การศึกษาแนวคิดการสร้างเครื่องมือวัด เป็นการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องเกี่ยวกับแนวคิดการสร้างเครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มี, 1 = มี เฉพาะการวิเคราะห์คุณภาพเครื่องมือ, 2 = มี เฉพาะวิธีการสร้างเครื่องมือ และ 3 = มี ทั้งวิธีการสร้างและการวิเคราะห์คุณภาพ

2.10 การสร้างเครื่องมือวัด เป็นตัวแปรเกี่ยวกับลักษณะการสร้างเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ประกอบด้วยรหัสตัวแปรดังนี้ 1 = ผู้วิจัยพัฒนาเครื่องมือวัดขึ้นเอง, 2 = ผู้วิจัยใช้เครื่องมือวัดจากงานวิจัยอื่น, 3 = ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยไทย, 4 = ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยต่างประเทศ และ 5 = ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยไทยและต่างประเทศ

2.11 การกำหนดนิยามเชิงปฏิบัติการ เป็นการกำหนดและนำเสนอเกี่ยวกับการกำหนดนิยามเชิงปฏิบัติการของตัวแปรที่สำคัญและองค์ประกอบที่กำหนดไว้ในเครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มี, 1 = มี แต่ไม่นำเสนอ และ 2 = มี และนำเสนอ

2.12 การระบุขั้นตอนในการพัฒนาเครื่องมือวัด เป็นการพิจารณาว่างานวิจัยได้ระบุขั้นตอนในการพัฒนาเครื่องมือวัดหรือไม่ ตั้งรหัสตัวแปรเป็นตัวแปรดัมมี่ 0 = ไม่มี และ 1 = มี

2.13 การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด เป็นการพิจารณาว่างานวิจัยได้กำหนดและนำเสนอตารางกำหนดโครงสร้างของเครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มี, 1 = มี แต่ไม่นำเสนอ และ 2 = มี และนำเสนอ

2.14 ประเภทของเครื่องมือวัด เป็นการระบุลักษณะของเครื่องมือวัดที่พัฒนาขึ้น ตั้งรหัสตัวแปรเป็นตัวแปรดัมมี่ 0 และ 1 โดยมีองค์ประกอบที่มุ่งวัด ดังนี้ แบบสอบหลายตัวเลือก, แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า, แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า, แบบสอบอัตนัย, แบบสอบภาคปฏิบัติ, แบบตรวจสอบรายการ, มาตรฐานค่า, แบบสังเกต, แบบสอบอัตนัยประยุกต์ (MEQ) และอื่น ๆ

2.15 จำนวนตัวเลือก/เกณฑ์ เป็นจำนวนตัวเลือก หรือเกณฑ์การให้คะแนนของเครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 1 = ไม่มีตัวเลือก, 2 = 2 ตัวเลือก, 3 = 3 ตัวเลือก, 4 = 4 ตัวเลือก, 5 = 5 ตัวเลือก, 6 = มากกว่า 5 ตัวเลือก และ 7 = ใช้มากกว่า 1 ประเภท

2.16 ลักษณะวิธีตอบของเครื่องมือวัด เป็นการระบุถึงลักษณะวิธีการตอบของเครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 1 = แบบเขียนตอบ (paper-pencil), 2 = คอมพิวเตอร์, 3 = ปฏิบัติจริง, 4 = ปากเปล่า และ 5 = อื่น ๆ

2.17 การตรวจให้คะแนน เป็นเกณฑ์การตรวจให้คะแนนของเครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 1 = ให้คะแนนแบบสองค่า, 2 = ให้คะแนนมากกว่าสองค่า, 3 = ให้คะแนนแบบสองค่า และมากกว่าสองค่า และ 9 = ไม่ระบุ

2.18 จำนวนครั้งที่ใช้ในการทดลอง เป็นการนำเครื่องมือวัดที่อยู่ระหว่างการพัฒนาไปทดลองใช้ (tryout) กับกลุ่มที่มีลักษณะคล้ายคลึงกับตัวอย่าง แต่ไม่ใช่ตัวอย่างในการวิจัยเพื่อตอบวัตถุประสงค์การวิจัย ประกอบด้วยรหัสตัวแปรดังนี้ 1 = 1 ครั้ง, 2 = 2 ครั้ง, 3 = 3 ครั้ง, 4 = 4 ครั้ง, 5 = 5 ครั้ง และ 9 = ไม่มี/ไม่ระบุ

2.19 ทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือวัด เป็นทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือวัดที่สร้างขึ้น ประกอบด้วยรหัสตัวแปรดังนี้ 1 = CTT, 2 = G Theory, 3 = IRT, 4 = CCT & G Theory, 5 = CCT & IRT, 6 = G Theory & IRT และ 7 = อื่น ๆ

2.20 วัตถุประสงค์ของการทดลองใช้ เป็นจุดมุ่งหมายของการนำเครื่องที่พัฒนาขึ้นไปทดลองใช้ (tryout) กับกลุ่มที่มีลักษณะคล้ายคลึงกับตัวอย่าง ตั้งรหัสตัวแปรเป็นตัวแปรดัมมี่ 0 และ 1 มีวัตถุประสงค์ของการทดลองใช้ ดังนี้ ตรวจสอบค่าความเที่ยง, ตรวจสอบค่าความตรง, ตรวจสอบความยาก, ตรวจสอบอำนาจจำแนก, ตรวจสอบภาษา, ตรวจสอบเวลาที่ใช้, ตรวจสอบตัวเลือก, ตรวจสอบความครอบคลุมของเนื้อหา ตรวจสอบคำชี้แจง/วิธีดำเนินการสอบ และทดสอบความไม่แปรเปลี่ยน

2.21 ดัชนีแสดงคุณภาพ เป็นการระบุเกณฑ์ประกอบการพิจารณาคุณภาพของเครื่องมือวัด ตั้งรหัสตัวแปรเป็นตัวแปรดัมมี่ ดังนี้ 0 = ไม่มี และ 1 = มี

2.22 ประเภทความเที่ยง เป็นการเลือกใช้ประเภทของการวิเคราะห์ความเที่ยงที่ใช้ในการตรวจสอบความคงเส้นคงวาของผลที่ได้จากการวัด ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มีการตรวจสอบ, 1 = test-retest method, 2 = measure of equivalence, 3 = test-retest with equivalent-forms method, 4 = split-half methods, 5 = KR20/KR21, 6 = Cronbach's alpha method, 7 = Lovett, 8 = Hoyt 's analysis, 9 = Kappa of Kohen, 10 = ความสอดคล้องระหว่างผู้ประเมิน, 11 = ใช้มากกว่า 1 วิธี, 12 = อื่น ๆ และ 99 = ไม่ระบุ

2.23 ประเภทความตรง เป็นการเลือกใช้ประเภทการตรวจสอบความตรงของเครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มีการตรวจสอบ, 1 = ความตรงตามเนื้อหา, 2 = ความตรงตามสภาพ, 3 = ความตรงเชิงทำนาย, 4 = ความตรงตามโครงสร้างด้วยการตัดสินโดยผู้เชี่ยวชาญ, 5 = ความตรงตามโครงสร้างด้วยวิธี known group, 6 = ความตรงตามโครงสร้างด้วยการวิเคราะห์เมทริกซ์พหุลักษณะวิธี, 7 = ความตรงตามโครงสร้างด้วยการวิเคราะห์องค์ประกอบ, 8 = ใช้มากกว่า 1 วิธี, 9 = อื่น ๆ และ 99 = ไม่ระบุ

2.24 วิธีการหาค่าความยาก เป็นการวิเคราะห์ความยากของข้อสอบ ประกอบด้วยรหัสตัวแปร ดังนี้ 0 = ไม่มีการตรวจสอบ, 1 = สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ, 2 = สูตร Whitney and Sabers, 3 = โปรแกรม Evana, 4 = โปรแกรม B-index, 5 = โปรแกรม TAP, 6 = IRT, 7 = อื่น ๆ, 8 = ใช้มากกว่า 1 วิธี และ 99 = ไม่ระบุ

2.25 วิธีการหาค่าอำนาจจำแนก เป็นการวิเคราะห์ค่าอำนาจจำแนกของข้อสอบ ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มีการตรวจสอบ, 1 = สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ, 2 = item-total correlation, 3 = สถิติทดสอบที (t-test), 4 = Biserial correlation, 5 = Point Biserial correlation, 6 = สูตร Whitney and Sabers, 7 = โปรแกรม Evana, 8 = โปรแกรม B-index, 9 = โปรแกรม TAP, 10 = IRT, 11 = อื่น ๆ 12 = มากกว่า 1 วิธี และ 99 = ไม่ระบุ

2.26 ประเภทของเกณฑ์ปกติ เป็นระดับของความสามารถการอ้างอิงโดยการนำคะแนนที่ได้ไปเทียบกับระดับของประชากรในกลุ่มนั้น ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มีการสร้างเกณฑ์ปกติ, 1 = เกณฑ์ปกติภายในชั้นเรียน, 2 = เกณฑ์ปกติระดับสถาบัน, 3 = เกณฑ์ปกติระดับท้องถิ่น, 4 = เกณฑ์ปกติระดับภาค, 5 = เกณฑ์ปกติระดับชาติ, 6 = เกณฑ์ปกติจำแนกตามระดับอายุ, 7 = เกณฑ์ปกติจำแนกตามระดับชั้นเรียน, 8 = อื่น ๆ และ 9 = ไม่ระบุ

2.27 การสร้างเกณฑ์ปกติ เป็นวิธีที่ใช้ในการสร้างเกณฑ์ปกติ ประกอบด้วยรหัสตัวแปรดังนี้ 0 = ไม่มีการสร้างเกณฑ์ปกติ, 2 = วิธีคะแนนมาตรฐานที่, 3 = วิธีคะแนนมาตรฐานที่ปกติ, 4 = วิธีสเตโนน, 5 =

วิธีเปอร์เซ็นต์ไทม์และคะแนนมาตรฐานที่รวมกัน, 6 = วิธีเปอร์เซ็นต์ไทม์และสเตรนจ์รวมกัน, 7 = วิธีคะแนนมาตรฐานที่ และจำแนกตามระดับชั้นเรียนรวมกัน และ 9 = ไม่ระบุ

2.28 คู่มือการใช้เครื่องมือวัด เป็นการระบุว่าคู่มือการใช้เครื่องมือวัดที่สร้างขึ้นหรือไม่ ตั้งรหัสตัวแปรเป็นตัวแปรตัวมี ดังนี้ 0 = ไม่มี และ 1 = มี

3. ตัวแปรคุณลักษณะของงานวิจัยด้านวิธีวิทยาการวิจัย ประกอบด้วย 7 ตัวแปร เป็นตัวแปรจัดประเภท 4 ตัวแปร ได้แก่ การกำหนดขนาดของตัวอย่าง, วิธีการเลือกตัวอย่าง, วิธีการเลือกตัวอย่างในการทดลองใช้ และการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย และมีตัวแปรต่อเนื่อง จำนวน 3 ตัวแปร ได้แก่ ขนาดของตัวอย่างในการวิจัย ขนาดของตัวอย่างในการทดลองใช้ และจำนวนผู้สอบในแต่ละฉบับ ทั้งนี้มีรายละเอียดของการกำหนดค่าของตัวแปรจัดประเภท ดังนี้

3.1 การกำหนดขนาดของตัวอย่าง เป็นวิธีที่ผู้วิจัยใช้ในการกำหนดขนาดของตัวอย่างในการวิจัย ประกอบด้วยรหัสตัวแปรดังนี้ 01= สูตรของ Krejcie & Morgan (1970), 02= สูตรของ Yamane (1973), 03= สูตรของ Darwin, 04= สูตรคำนวณของบุญชม ศรีสะอาด (2553), 05= เปิดตารางของ ศิริชัย กาญจนวาสิ และคณะ (2535), 06= กำหนดขนาดตัวอย่างตามสัดส่วน, 07= เกณฑ์ร้อยละของประชากร, 08= สูตรกำหนดขนาดด้วยการสุ่มแบบแบ่งชั้น, 09= เกณฑ์กำหนดขนาดตัวอย่างสำหรับพัฒนาเครื่องมือ ของ บุญชม ศรีสะอาด (2534), 10= โปรแกรม G* Power, 11= 10 เท่าของตัวแปร/พารามิเตอร์, 12= 15 เท่าของตัวแปร/พารามิเตอร์, 13= 20 เท่าของตัวแปร/พารามิเตอร์, 14= หลาย ๆ เกณฑ์, 15= อื่น ๆ และ 99= ไม่ระบุ

3.2 วิธีการเลือกตัวอย่าง เป็นวิธีที่ผู้วิจัยใช้เพื่อการได้มาซึ่งตัวอย่างในการวิจัย ประกอบด้วยรหัสตัวแปรดังนี้ 1 = ใช้ประชากร/กลุ่มเป้าหมาย, 2 = สุ่มอย่างง่าย, 3 = สุ่มแบบกลุ่ม, 4 = สุ่มแบบแบ่งชั้น, 5 = สุ่มแบบสองขั้นตอน, 6 = สุ่มแบบหลายขั้นตอน, 7 = เลือกอย่างเจาะจง, 8 = อื่น ๆ และ 9 = ไม่ระบุ

3.3 วิธีการเลือกตัวอย่างในการทดลองใช้ เป็นวิธีที่ผู้วิจัยใช้เพื่อการได้มาซึ่งตัวอย่างในการทดลองใช้เครื่องมือวัด ประกอบด้วยรหัสตัวแปรดังนี้ 1 = ใช้ประชากร/กลุ่มเป้าหมาย, 2 = สุ่มอย่างง่าย, 3 = สุ่มแบบกลุ่ม, 4 = สุ่มแบบแบ่งชั้น, 5 = สุ่มแบบสองขั้นตอน, 6 = สุ่มแบบหลายขั้นตอน, 7 = เลือกอย่างเจาะจง, 8 = อื่น ๆ และ 9 = ไม่ระบุ

3.4 การวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย เป็นสถิติที่ใช้ในการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย ประกอบด้วยรหัสตัวแปรดังนี้ content analysis, descriptive stat, independent t-test, dependent t-test, chi-square, correlation, regression, one-way ANOVA, two-way ANOVA, ANCOVA, MANOVA, exploratory factor analysis, confirmatory factor analysis, ทดสอบความไม่แปรเปลี่ยน, ค่าความเที่ยง, ค่าความตรง, ค่าความยาก และค่าอำนาจจำแนก

4. ตัวแปรคุณลักษณะของงานวิจัยด้านคุณภาพงานวิจัย มี 1 ตัวแปร เป็นตัวแปรต่อเนื่อง คือ คะแนนประเมินคุณภาพงานวิจัย

จากตัวแปรที่ใช้ในการวิจัยดังกล่าวข้างต้น สามารถจำแนกเป็นตัวแปรคุณลักษณะงานวิจัยได้ 2 ส่วนคือ ตัวแปรคุณลักษณะงานวิจัยที่ใช้สำหรับการวิเคราะห์ข้อมูลพื้นฐานเพื่อบรรยายให้เห็นถึงลักษณะของงานวิจัยที่นำมาสังเคราะห์ และตัวแปรคุณลักษณะของงานวิจัยสำหรับศึกษาผลที่มีต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ตามที่ได้กำหนดไว้ในภาพที่ 2.1 กรอบแนวคิดในการวิจัย สามารถสรุปรายละเอียดตัวแปรคุณลักษณะงานวิจัยจำแนกตามการนำไปใช้วิเคราะห์ข้อมูล รายละเอียดดังตารางที่ 3.1

ตารางที่ 3.1 รายละเอียดตัวแปรคุณลักษณะงานวิจัยจำแนกตามการนำไปใช้วิเคราะห์ข้อมูล

คุณลักษณะของงานวิจัย	ตัวแปร	ข้อมูลพื้นฐาน	การทดสอบ ตัวแปร คุณลักษณะ*	
1. ด้านการพิมพ์และผู้วิจัย				
ตัวแปรจัดประเภท	1. ปีที่พิมพ์งานวิจัย	✓	✓	
	2. ประเภทสถาบันที่ผลิตงานวิจัย	✓	✓	
	3. สถาบันที่ผลิตงานวิจัย	✓	✓	
	4. สาขาที่ผลิตงานวิจัย	✓	✓	
	5. ประเภทงานวิจัย	✓	✓	
	6. เพศของผู้วิจัย	✓	✓	
2. ด้านเนื้อหาสาระงานวิจัย				
ตัวแปรจัดประเภท	1. วัตถุประสงค์ของการวิจัย	✓	-	
	2. เครื่องมือวัดที่พัฒนาขึ้น	✓	-	
	3. แนวคิด/ทฤษฎีหลัก	✓	-	
	4. องค์ประกอบที่มุ่งวัด	✓	✓	
	5. เนื้อหาที่เกี่ยวข้อง	✓	✓	
	6. ภูมิภาค	✓	✓	
	7. แหล่งที่มาของตัวอย่าง	✓	✓	
	8. ระดับชั้นของตัวอย่าง	✓	✓	
	9. การศึกษาแนวคิดการสร้างเครื่องมือวัด	✓	✓	
	10. การสร้างเครื่องมือวัด	✓	✓	
	11. การกำหนดนิยามเชิงปฏิบัติการ	✓	✓	
	12. การระบุขั้นตอนในการพัฒนาเครื่องมือวัด	✓	-	
	13. การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด	✓	✓	
	14. ประเภทของเครื่องมือวัด	✓	✓	
	15. จำนวนตัวเลือก/เกณฑ์	✓	✓	
	16. ลักษณะวิธีตอบของเครื่องมือวัด	✓	✓	
	17. การตรวจให้คะแนน	✓	✓	
	18. จำนวนครั้งที่ใช้ในการทดลอง	✓	✓	
	19. ทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือวัด	✓	-	
	20. วัตถุประสงค์ของการทดลองใช้	✓	-	
	21. ดัชนีแสดงคุณภาพ	✓	-	
	22. ประเภทความเที่ยง	✓	-	
	23. ประเภทความตรง	✓	-	
	24. วิธีการหาค่าความยาก	✓	-	
	25. วิธีการหาค่าอำนาจจำแนก	✓	-	
	26. ประเภทของเกณฑ์ปกติ	✓	-	
	27. วิธีที่ใช้ในการสร้างเกณฑ์ปกติ	✓	✓	
	28. คู่มือการใช้เครื่องมือวัด	✓	✓	
	ตัวแปรต่อเนื่อง	29. จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา	✓	✓
		30. จำนวนผู้เชี่ยวชาญด้านการวัดผล	✓	-
		31. จำนวนผู้เชี่ยวชาญด้านการเนื้อหา	✓	-

คุณลักษณะของงานวิจัย	ตัวแปร	ข้อมูลพื้นฐาน	การทดสอบ ตัวแปร คุณลักษณะ*
	32. จำนวนข้อที่สร้างก่อนการทดลองใช้ (รวมทั้งหมด)	✓	-
	33. จำนวนข้อคำถามใช้จริง (รวม)	✓	-
	34. จำนวนข้อคำถามในแต่ละฉบับ	✓	✓
3. ด้านวิธีวิทยาการวิจัย			
ตัวแปรจัดประเภท	1. การกำหนดขนาดของตัวอย่าง	✓	-
	2. วิธีการเลือกตัวอย่าง	✓	✓
	3. วิธีการเลือกตัวอย่างในการทดลองใช้	✓	-
	4. การวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย	✓	-
ตัวแปรต่อเนื่อง	5. ขนาดของตัวอย่างในการวิจัย	✓	-
	6. ขนาดของตัวอย่างในการทดลองใช้	✓	-
	7. จำนวนผู้สอบในแต่ละฉบับ	✓	✓
4. ด้านคุณภาพงานวิจัย			
ตัวแปรต่อเนื่อง	คะแนนประเมินคุณภาพงานวิจัย	✓	✓

* ผู้วิจัยมีการปรับให้เป็นตัวแปรจัดประเภททั้งหมด

เครื่องมือที่ใช้ในการวิจัย

การวิจัยครั้งนี้ประกอบด้วยเครื่องมือที่ใช้ในการวิจัย 2 ฉบับ คือ แบบประเมินคุณภาพงานวิจัย และแบบบันทึกคุณลักษณะงานวิจัย ได้แก่

1. แบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพงานวิจัย
2. แบบบันทึกคุณลักษณะงานวิจัย และคู่มือลงรหัส (code book) สำหรับบันทึกข้อมูลคุณลักษณะงานวิจัย

วิธีการสร้างและการตรวจสอบคุณภาพ

1. แบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพงานวิจัย

แบบประเมินคุณภาพงานวิจัยเป็นเครื่องมือสำหรับประเมินคุณภาพงานวิจัย พร้อมเกณฑ์การประเมิน สำหรับใช้ประเมินงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ขั้นตอนการสร้างแบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพงานวิจัย มีรายละเอียดดังนี้

1.1 ผู้วิจัยพัฒนาแบบประเมินคุณภาพงานวิจัย โดยปรับปรุงจากแบบประเมินคุณภาพงานวิจัยของสำนักเลขาธิการสภาการศึกษา (2552) ไชยยศ ไพวิทยศิริธรรม (2544) อาภาพรรณ ประทุมไทย (2553) และสุวิมล เสวกสุริยวงศ์ (2553) แบบประเมินคุณภาพงานวิจัยที่พัฒนาขึ้นมีลักษณะเป็นมาตราประมาณค่า 5 ระดับ โดยกำหนดเกณฑ์การให้คะแนนแต่ละระดับ (rubric) ตั้งแต่ 0-4 มีเนื้อหาสาระในการประเมิน 6 ด้าน ได้แก่ ความเป็นมาและความสำคัญของปัญหาวิจัย เอกสารและงานวิจัยที่เกี่ยวข้อง วิธีดำเนินการวิจัย ผลการวิเคราะห์ข้อมูล การสรุปอภิปรายผลและข้อเสนอแนะ และการนำเสนอรายงานวิจัย จำนวน 34 ข้อ มีโครงสร้างของแบบประเมินคุณภาพงานวิจัยดังตารางที่ 3.2

ตารางที่ 3.2 โครงสร้างของแบบประเมินคุณภาพงานวิจัย

เนื้อหาสาระที่ประเมิน	จำนวนข้อ	ร้อยละ	ข้อที่
1. ความเป็นมาและความสำคัญของปัญหาวิจัย	7	20.59	1-7
2. เอกสารและงานวิจัยที่เกี่ยวข้อง	4	11.76	8-11
3. วิธีดำเนินการวิจัย			
3.1 ระเบียบวิธีวิจัย	2	5.88	12-13
3.2 เกณฑ์การเลือกกลุ่มตัวอย่าง	2	5.88	14-15
3.3 เครื่องมือและการเก็บรวบรวมข้อมูล	9	26.47	16-24
3.4 การออกแบบการวิเคราะห์ข้อมูล	2	5.88	25-26
4. ผลการวิเคราะห์ข้อมูล	2	5.88	27-28
5. การสรุปอภิปรายผลและข้อเสนอแนะ	3	8.82	29-31
6. การนำเสนอรายงานวิจัย	3	8.82	32-34
รวม	34	100	

โดยมีเกณฑ์ในการแปลผลการประเมินคุณภาพงานวิจัย แบ่งออกเป็นระดับคุณภาพ 5 ระดับ 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

1.2 ผู้วิจัยนำแบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพงานวิจัยที่พัฒนาขึ้นเสนอต่อผู้เชี่ยวชาญ จำนวน 5 ท่าน ซึ่งมีความเชี่ยวชาญในเรื่องของการวิเคราะห์ห่อภิมาณ การวัดประเมินผลทางการศึกษา และวิจัยทางการศึกษา เพื่อตรวจสอบความตรงเชิงเนื้อหา (content validity) ผลการวิเคราะห์ค่าความสอดคล้องระหว่างเนื้อหาสาระที่ประเมินและเกณฑ์การประเมินคุณภาพงานวิจัย (IOC: index of item objective congruence) มีค่าอยู่ระหว่าง 0.80-1.00 (รายละเอียดดังแสดงในภาคผนวก ข) อย่างไรก็ตามผู้เชี่ยวชาญได้ให้ข้อเสนอแนะสำหรับปรับปรุงเนื้อหาสาระที่ประเมินและเกณฑ์การประเมินคุณภาพผู้วิจัยได้ดำเนินการปรับตามข้อเสนอแนะมีรายละเอียดดังตารางที่ 3.3

ตารางที่ 3.3 ผลการปรับปรุงแบบประเมินคุณภาพงานวิจัย และเกณฑ์การประเมินคุณภาพตามข้อเสนอแนะของผู้เชี่ยวชาญ

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
1) เสนอแนะให้ปรับนิยามการให้คะแนนเป็น 1 หมายถึง มีการให้นิยามตัวแปรสำคัญที่ศึกษาอย่างครบถ้วนแต่เป็นนิยามเชิงทฤษฎี	ผู้วิจัยปรับภาษาตามข้อเสนอแนะ ดังนี้
2) เปลี่ยนจากคำว่า “องค์ประกอบ” เป็น “ตัวแปรที่สำคัญ”	
6. การนิยามศัพท์เฉพาะมีความชัดเจน 0 หมายถึง ไม่มีการนิยามศัพท์เฉพาะ หรือมีการให้นิยามเชิงทฤษฎีแต่ไม่ครบถ้วน	6. การนิยามศัพท์เฉพาะมีความชัดเจน 0 หมายถึง ไม่มีการนิยามศัพท์เฉพาะ หรือมีการให้นิยามเชิงทฤษฎีแต่ไม่ครบถ้วน

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
<p>1 หมายถึง มีการให้นิยามเชิงทฤษฎี เฉพาะตัวแปรสำคัญที่ศึกษาอย่างครบถ้วน</p> <p>2 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัด แต่ไม่ครบทุกองค์ประกอบ</p> <p>3 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัด ครบทุกองค์ประกอบ</p> <p>4 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัด ครบทุกองค์ประกอบ โดยบ่งบอกลักษณะที่สำคัญในรูปที่สามารถสังเกตและวัดได้</p>	<p>1 หมายถึง มีการให้นิยามตัวแปรสำคัญที่ศึกษาอย่างครบถ้วนแต่เป็นนิยามเชิงทฤษฎี</p> <p>2 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัด แต่ไม่ครบทุกตัวแปรสำคัญที่ศึกษา</p> <p>3 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัด ครบทุกตัวแปรสำคัญที่ศึกษา</p> <p>4 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัด ครบทุกตัวแปรสำคัญที่ศึกษา โดยบ่งบอกลักษณะที่สำคัญในรูปที่สามารถสังเกตและวัดได้</p>
<p>มีข้อสังเกตว่า หากผลการพิจารณาเป็น 30% พอได้คะแนนเท่าไร</p> <p>10. เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัยและเชื่อมโยงกับงานวิจัยในอดีต</p> <p>0 หมายถึง ไม่มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปีเลย</p> <p>1 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณไม่ถึง 30%</p> <p>2 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณมากกว่า 30% แต่ไม่ถึง 50%</p> <p>3 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณมากกว่า 50% แต่ไม่ถึง 70%</p> <p>4 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณมากกว่า 70% ขึ้นไป</p>	<p>ผู้วิจัยปรับภาษาที่สะท้อนช่วงการพิจารณาให้ชัดเจนยิ่งขึ้น ดังนี้</p> <p>10. เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัยและเชื่อมโยงกับงานวิจัยในอดีต</p> <p>0 หมายถึง ไม่มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปีเลย</p> <p>1 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณไม่ถึง 30%</p> <p>2 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณตั้งแต่ 30% แต่ไม่ถึง 50%</p> <p>3 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณตั้งแต่ 50% แต่ไม่ถึง 70%</p> <p>4 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณตั้งแต่ 70% ขึ้นไป</p>
<p>หากสัดส่วนไม่ได้เป็นไปตามที่กำหนดพอที่จะพิจารณาอย่างไร</p> <p>11. เอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศมีสัดส่วนที่เหมาะสม</p> <p>0 หมายถึง ไม่มีเอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศ</p> <p>1 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีเพียงภาษาไทยเท่านั้น</p> <p>2 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทย 80% และภาษาต่างประเทศ 20%</p>	<p>ผู้วิจัยปรับภาษาที่สะท้อนช่วงการพิจารณาให้ชัดเจนยิ่งขึ้น ดังนี้</p> <p>11. เอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศมีสัดส่วนที่เหมาะสม</p> <p>0 หมายถึง ไม่มีเอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศ</p> <p>1 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีเพียงภาษาไทยเท่านั้น</p> <p>2 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทยมากกว่า 80% และภาษาต่างประเทศน้อยกว่า 20%</p>

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
<p>3 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทย 50% และภาษาต่างประเทศ 50%</p> <p>4 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทย 30% และภาษาต่างประเทศ 70%</p>	<p>3 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทย 50% และภาษาต่างประเทศ 50%</p> <p>4 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทยน้อยกว่า 30% และภาษาต่างประเทศมากกว่า 70%</p>
<p>ใช้คำว่า “ประชากร” และ “ตัวอย่างวิจัย”</p> <p>14. กำหนดประชากรและกลุ่มตัวอย่างได้อย่างเหมาะสมกับการวิจัย</p> <p>0 หมายถึง ไม่ระบุกลุ่มประชากร และไม่ระบุกลุ่มตัวอย่าง</p> <p>1 หมายถึง ไม่ระบุกลุ่มประชากร แต่ระบุกลุ่มตัวอย่าง</p> <p>2 หมายถึง ระบุกลุ่มประชากร ระบุกลุ่มตัวอย่าง แต่ไม่มีการสุ่มตัวอย่าง</p> <p>3 หมายถึง ระบุกลุ่มประชากร ระบุกลุ่มตัวอย่าง ระบุการสุ่มตัวอย่าง แต่ให้รายละเอียดในการสุ่มกลุ่มตัวอย่างไม่ชัดเจน</p> <p>4 หมายถึง ระบุกลุ่มประชากร ระบุกลุ่มตัวอย่าง ระบุการสุ่มตัวอย่าง ให้รายละเอียดในการสุ่มกลุ่มตัวอย่างชัดเจน อ่านเข้าใจง่าย</p>	<p>ผู้วิจัยปรับภาษาตามข้อเสนอแนะ ดังนี้</p> <p>14. กำหนดประชากรและตัวอย่างได้อย่างเหมาะสมกับการวิจัย</p> <p>0 หมายถึง ไม่ระบุประชากร และตัวอย่าง</p> <p>1 หมายถึง ระบุประชากร และตัวอย่างเพียงอย่างใดอย่างหนึ่ง</p> <p>2 หมายถึง ระบุประชากร ระบุตัวอย่าง แต่ไม่มีการสุ่มตัวอย่าง</p> <p>3 หมายถึง ระบุประชากร ระบุตัวอย่าง ระบุการสุ่มตัวอย่าง แต่ให้รายละเอียดในการสุ่มตัวอย่างไม่ชัดเจน</p> <p>4 หมายถึง ระบุประชากร ระบุตัวอย่าง ระบุการสุ่มตัวอย่าง ให้รายละเอียดในการสุ่มตัวอย่างชัดเจน อ่านเข้าใจง่าย</p>
<p>1) ใช้คำว่า “ตัวอย่างวิจัย”</p> <p>2) การเลือกอย่างเจาะจง ในเกณฑ์การให้ 1 คะแนน เป็นส่วนหนึ่งของการเลือกกลุ่มตัวอย่างโดยไม่อาศัยความน่าจะเป็น ในเกณฑ์การให้ 2 คะแนน</p> <p>3) เพิ่มใช้ประชากรในการศึกษา</p> <p>15. ขนาดกลุ่มตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม</p> <p>0 หมายถึง ไม่มีการระบุการได้มาของกลุ่มตัวอย่าง และขนาดกลุ่มตัวอย่าง</p> <p>1 หมายถึง ระบุขนาดกลุ่มตัวอย่าง และกลุ่มตัวอย่างได้มาจากการเลือกอย่างเจาะจง</p> <p>2 หมายถึง ระบุขนาดกลุ่มตัวอย่าง และกลุ่มตัวอย่างได้มาจากการเลือกกลุ่มตัวอย่างโดยไม่อาศัยความน่าจะเป็น</p> <p>3 หมายถึง ระบุขนาดกลุ่มตัวอย่าง และกลุ่มตัวอย่างได้มาจากการสุ่มโดยอาศัยความน่าจะเป็น และเป็นตัวแทนที่ดีของประชากร แต่มีที่มาของการกำหนดขนาดของกลุ่มตัวอย่างไม่สมเหตุสมผล</p>	<p>ผู้วิจัยปรับภาษา และเกณฑ์การตามข้อเสนอแนะ ดังนี้</p> <p>15. ขนาดตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม</p> <p>0 หมายถึง ไม่มีการระบุการได้มาของตัวอย่างและขนาดตัวอย่าง หรือศึกษากับประชากร</p> <p>1 หมายถึง ระบุขนาดตัวอย่าง แต่ไม่มีการระบุการได้มาของตัวอย่าง</p> <p>2 หมายถึง ระบุขนาดตัวอย่าง และตัวอย่างได้มาจากการเลือกตัวอย่างโดยไม่อาศัยความน่าจะเป็น</p> <p>3 หมายถึง ระบุขนาดตัวอย่าง และตัวอย่างได้มาจากการสุ่มโดยอาศัยความน่าจะเป็น และเป็นตัวแทนที่ดีของประชากร แต่มีที่มาของการกำหนดขนาดของตัวอย่างไม่สมเหตุสมผล</p>

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
<p>4 หมายถึง ระบุขนาดกลุ่มตัวอย่าง และกลุ่มตัวอย่างได้มาจากการสุ่มโดยอาศัยความน่าจะเป็น เป็นตัวแทนที่ดีของประชากร และมีที่มาของการกำหนดขนาดของกลุ่มตัวอย่างสมเหตุสมผล</p>	<p>4 หมายถึง ระบุขนาดตัวอย่าง และตัวอย่างได้มาจากการสุ่มโดยอาศัยความน่าจะเป็น เป็นตัวแทนที่ดีของประชากร และมีที่มาของการกำหนดขนาดของกลุ่มตัวอย่างสมเหตุสมผล</p>
<p>ปรับเกณฑ์การให้ 4 คะแนน เป็น “มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน ครบถ้วน และสอดคล้องกับความต้องการของสังคมยุคปัจจุบันในการนำไปใช้ประโยชน์”</p> <p>16. การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน</p> <p>0 หมายถึง ไม่มีการกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21</p> <p>1 หมายถึง มีการกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 แต่ไม่ชัดเจน</p> <p>2 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน แต่ไม่สอดคล้องกับความต้องการนำไปใช้ประโยชน์</p> <p>3 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน และสอดคล้องกับความต้องการนำไปใช้ประโยชน์ แต่ไม่ครบถ้วนในบางประเด็น</p> <p>4 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน ครบถ้วน และสอดคล้องกับความต้องการนำไปใช้ประโยชน์</p>	<p>ผู้วิจัยปรับเกณฑ์การตามข้อเสนอแนะ ดังนี้</p> <p>16. การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน</p> <p>0 หมายถึง ไม่มีการกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21</p> <p>1 หมายถึง มีการกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 แต่ไม่ชัดเจน</p> <p>2 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน แต่ไม่สอดคล้องกับความต้องการนำไปใช้ประโยชน์</p> <p>3 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน และสอดคล้องกับความต้องการนำไปใช้ประโยชน์ แต่ไม่ครบถ้วนในบางประเด็น</p> <p>4 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน ครบถ้วน และสอดคล้องกับความต้องการของสังคมยุคปัจจุบันในการนำไปใช้ประโยชน์</p>
<p>ปรับคำว่า “วิธี” เป็น “ขั้นตอน”</p> <p>20. วิธีในการดำเนินการวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม</p> <p>0 หมายถึง ขั้นตอนในการดำเนินการวัดไม่ชัดเจน ไม่มีการระบุค่าชี้แจงในเครื่องมือ</p> <p>1 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน แต่ไม่มีการระบุค่าชี้แจงในเครื่องมือ</p> <p>2 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียด แต่ไม่ครบถ้วน</p> <p>3 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียดครบถ้วน</p>	<p>ผู้วิจัยปรับภาษาตามข้อเสนอแนะ ดังนี้</p> <p>20. ขั้นตอนในการดำเนินการวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม</p> <p>0 หมายถึง ขั้นตอนในการดำเนินการวัดไม่ชัดเจน ไม่มีการระบุค่าชี้แจงในเครื่องมือ</p> <p>1 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน แต่ไม่มีการระบุค่าชี้แจงในเครื่องมือ</p> <p>2 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียด แต่ไม่ครบถ้วน</p> <p>3 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียดครบถ้วน</p>

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
<p>4 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียดครบถ้วน พร้อมยกตัวอย่างวิธีการทำ</p>	<p>4 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียดครบถ้วน พร้อมยกตัวอย่างวิธีการทำ</p>
<p>เปลี่ยนจากคำว่า “แบบวัด” เป็น “เครื่องมือวัด”</p> <p>23. กลุ่มตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม</p> <p>0 หมายถึง ไม่มีการทดลองใช้แบบวัด</p> <p>1 หมายถึง กลุ่มตัวอย่างที่ทดลองใช้แบบวัดเป็นส่วนหนึ่งของกลุ่มตัวอย่างที่จะใช้จริง</p> <p>2 หมายถึง กลุ่มตัวอย่างที่ทดลองใช้แบบวัดสุ่มมาจากประชากรกลุ่มเดียวกัน แต่ไม่ได้ระบุจำนวนของกลุ่มตัวอย่าง</p> <p>3 หมายถึง กลุ่มตัวอย่างที่ทดลองใช้แบบวัดสุ่มมาจากประชากรกลุ่มเดียวกัน แต่จำนวนที่ใช้ไม่พอเพียง/พอเหมาะ</p> <p>4 หมายถึง กลุ่มตัวอย่างที่ทดลองใช้แบบวัดสุ่มมาจากประชากรกลุ่มเดียวกัน และจำนวนที่ใช้พอเพียง/พอเหมาะ</p>	<p>ผู้วิจัยปรับภาษาตามข้อเสนอแนะ ดังนี้</p> <p>23. ตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม</p> <p>0 หมายถึง ไม่มีการทดลองใช้เครื่องมือวัด</p> <p>1 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดเป็นส่วนหนึ่งของตัวอย่างที่จะใช้จริง</p> <p>2 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดสุ่มมาจากประชากรกลุ่มเดียวกัน แต่ไม่ได้ระบุจำนวนของตัวอย่าง</p> <p>3 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดสุ่มมาจากประชากรกลุ่มเดียวกัน แต่จำนวนที่ใช้ไม่พอเพียง/พอเหมาะ</p> <p>4 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดสุ่มมาจากประชากรกลุ่มเดียวกัน และจำนวนที่ใช้พอเพียง/พอเหมาะ</p>
<p>คุณภาพของเครื่องมือในภาพรวมพิจารณาอย่างไร ควรพิจารณา ค่าความยาก ความเที่ยง อำนาจจำแนก ให้ชัดเจน</p> <p>24. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ในภาพรวมมีคุณภาพเหมาะสม</p> <p>0 หมายถึง คุณภาพของแบบวัดในภาพรวมอยู่ในระดับต่ำมาก</p> <p>1 หมายถึง คุณภาพของแบบวัดในภาพรวมอยู่ในระดับต่ำ</p> <p>2 หมายถึง คุณภาพของแบบวัดในภาพรวมอยู่ในระดับปานกลาง</p> <p>3 หมายถึง คุณภาพของแบบวัดในภาพรวมอยู่ในระดับดี</p> <p>4 หมายถึง คุณภาพของแบบวัดในภาพรวมอยู่ในระดับดีมาก</p>	<p>ผู้วิจัยปรับเกณฑ์การตามข้อเสนอแนะ ดังนี้</p> <p>24. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ในภาพรวมมีคุณภาพเหมาะสม</p> <p>0 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับต่ำมาก (ค่าความเที่ยงไม่ถึง 0.7 และไม่รายงานค่าความยาก อำนาจจำแนก/ไม่เป็นไปตามเกณฑ์)</p> <p>1 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับต่ำ (ค่าความเที่ยงไม่ถึง 0.7 ค่าความยาก และอำนาจจำแนกเป็นไปตามเกณฑ์)</p> <p>2 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับปานกลาง (ค่าความเที่ยง 0.7 ขึ้นไป ค่าความยาก และอำนาจจำแนกเป็นไปตามเกณฑ์)</p> <p>3 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับดี (ค่าความเที่ยง 0.8 ขึ้นไป ค่าความยาก และอำนาจจำแนกเป็นไปตามเกณฑ์)</p>

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
	4 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับดีมาก (ค่าความเที่ยง 0.9 ขึ้นไป ค่าความยาก และอำนาจจำแนก เป็นไปตามเกณฑ์)
น่าจะมีการปรับว่า 5 ส่วนมีถูกกี่ส่วน เช่น มีถูกเพียง 1 ส่วนหรือไม่ถูกเลย = 0, ถูก 2 ส่วน = 1 เป็นต้น	ผู้วิจัยปรับเกณฑ์การตามข้อเสนอแนะ ดังนี้
34. รูปแบบรายงานถูกต้องตามหลักวิชา	34. รูปแบบรายงานถูกต้องตามหลักวิชา (5 ส่วน คือ บทนำ รายงานเอกสารที่เกี่ยวข้อง วิธีการวิจัย ผลการวิเคราะห์ข้อมูล และสรุปและอภิปรายผล)
0 หมายถึง รูปแบบรายงานถูกต้องเพียง 1 ส่วน คือ บทนำ	0 หมายถึง รูปแบบรายงานถูกต้องเพียง 1 ส่วน
1 หมายถึง รูปแบบรายงานถูกต้อง 2 ส่วน คือ บทนำ และรายงานเอกสารที่เกี่ยวข้อง	1 หมายถึง รูปแบบรายงานถูกต้อง 2 ส่วน
2 หมายถึง รูปแบบรายงานถูกต้อง 3 ส่วน คือ บทนำ รายงานเอกสารที่เกี่ยวข้อง และวิธีการวิจัย	2 หมายถึง รูปแบบรายงานถูกต้อง 3 ส่วน
3 หมายถึง รูปแบบรายงานถูกต้อง 4 ส่วน คือ บทนำ รายงานเอกสารที่เกี่ยวข้อง วิธีการวิจัย และผลการวิเคราะห์ข้อมูล	3 หมายถึง รูปแบบรายงานถูกต้อง 4 ส่วน
4 หมายถึง รูปแบบรายงานถูกต้อง 5 ส่วน คือ บทนำ รายงานเอกสารที่เกี่ยวข้อง วิธีการวิจัย ผลการวิเคราะห์ข้อมูล และสรุปและอภิปรายผล	4 หมายถึง รูปแบบรายงานถูกต้อง 5 ส่วน

1.3 ผู้วิจัยได้นำแบบประเมินคุณภาพงานวิจัยและเกณฑ์การประเมินคุณภาพงานวิจัยที่แก้ไขและปรับปรุงตามคำแนะนำของผู้เชี่ยวชาญ มาตรวจสอบความสอดคล้องในการประเมินคุณภาพงานวิจัย (inter rater reliability) โดยคณะผู้วิจัยทั้ง 2 คน ศึกษาทำความเข้าใจเกณฑ์การให้คะแนนแต่ละระดับอย่างละเอียด จากนั้นทดลองฝึกประเมินงานวิจัยจำนวน 3 เล่ม ที่ได้จากการสุ่ม โดยสุ่มจากงานวิจัยจากมหาวิทยาลัยในกำกับ มหาวิทยาลัยราชภัฏ และสถาบันผู้ให้ทุน มากลุ่มละ 1 เล่ม ผู้วิจัยทั้ง 2 คน ประเมินงานวิจัยทีละเล่ม นำผลการประเมินของแต่ละคนมาพูดคุยเพื่อสร้างความเข้าใจให้ตรงกันก่อนมีการประเมินเล่มต่อไป และตรวจสอบความสอดคล้องในการประเมินคุณภาพงานวิจัย ตามสูตรของ Copper & Hedges (1994 อ้างถึงใน วัชญา ยิ้มยวน, 2547) โดยพิจารณาจากสัดส่วนของจำนวนข้อที่ในการประเมินงานวิจัยที่สอดคล้องกันกับจำนวนข้อการประเมินงานวิจัยทั้งหมด นอกจากนี้ ผู้วิจัยได้ดำเนินการตรวจสอบความสอดคล้องภายในตัวผู้ประเมิน (intra-rater reliability) กล่าวคือ ผู้วิจัยทั้ง 2 คน ประเมินคุณภาพงานวิจัยทั้ง 3 เล่ม ซ้ำอีกครั้งหนึ่ง โดยทิ้งระยะห่างจากการประเมินครั้งแรก 2 สัปดาห์ และนำข้อมูลที่ได้จากการประเมินทั้ง 2 ครั้งมาตรวจสอบความสอดคล้องในการให้คะแนนตามสูตรดังนี้

สูตรในการคำนวณค่าความสอดคล้องในการประเมิน

$$\text{ค่าความสอดคล้องในการประเมิน (Agreement Rate: AR)} = \frac{\text{จำนวนข้อในการประเมินคุณภาพงานวิจัยที่สอดคล้องกัน}}{\text{จำนวนข้อของการประเมินคุณภาพงานวิจัยทั้งหมด}}$$

เกณฑ์ในการแปลผลค่าความสอดคล้องในการประเมิน

AR มีค่ามากกว่า 0.75	หมายถึง สอดคล้องกันดีมาก
AR มีค่าอยู่ระหว่าง 0.60 ถึง 0.74	หมายถึง สอดคล้องกันดี
AR มีค่าอยู่ระหว่าง 0.40 ถึง 0.59	หมายถึง สอดคล้องกันปานกลาง
AR มีค่าน้อยกว่า 0.40	หมายถึง ไม่สอดคล้องกัน

ผลการตรวจสอบความสอดคล้องระหว่างผู้ประเมิน พบว่า มีค่าความสอดคล้องอยู่ระหว่าง 0.85-0.91 สรุปได้ว่า ผู้วิจัยทั้ง 2 คน สามารถประเมินคุณภาพงานวิจัยได้สอดคล้องใกล้เคียงกันอยู่ในระดับดีมาก ในขณะที่ผลการตรวจสอบความสอดคล้องภายในตัวผู้ประเมิน พบว่า ผู้วิจัยคนที่ 1 มีค่าความสอดคล้องอยู่ระหว่าง 0.88-0.94 และผู้วิจัยคนที่ 2 มีค่าความสอดคล้องอยู่ระหว่าง 0.88-0.97 จึงสรุปได้ว่า ผู้วิจัยแต่ละคนสามารถประเมินคุณภาพงานวิจัยได้ผลการประเมินที่คงที่อยู่ในระดับดีมาก รายละเอียดดังตารางที่ 3.4

ตารางที่ 3.4 ผลการตรวจสอบความสอดคล้องระหว่างผู้ประเมินและภายในตัวผู้ประเมินจากการประเมินคุณภาพงานวิจัย

เล่มงานวิจัย	ผู้วิจัยคนที่ 1 และ 2		ผู้วิจัยคนที่ 1 ประเมินครั้งที่ 2		ผู้วิจัยคนที่ 2 ประเมินครั้งที่ 2	
	จำนวนข้อที่ประเมินตรงกัน	AR	จำนวนข้อที่ประเมินตรงกัน	AR	จำนวนข้อที่ประเมินตรงกัน	AR
เล่มที่ 1	29	0.85	30	0.88	30	0.88
เล่มที่ 2	31	0.91	32	0.94	31	0.91
เล่มที่ 3	31	0.91	32	0.94	33	0.97

(AR: Agreement Rate, ค่าความสอดคล้องในการประเมิน)

2. แบบบันทึกคุณลักษณะงานวิจัย และคู่มือลงรหัส

แบบบันทึกคุณลักษณะงานวิจัยและคู่มือลงรหัส เป็นเครื่องมือสำหรับบันทึกคุณลักษณะต่าง ๆ ของงานวิจัยตามที่คุณวิจัยได้กำหนดไว้ แบ่งออกเป็น 4 ด้าน ได้แก่ ด้านการพิมพ์และผู้วิจัย ด้านเนื้อหาสาระงานวิจัย ด้านวิธีวิทยาการวิจัย และด้านคุณภาพงานวิจัย ขั้นตอนการสร้างแบบบันทึกคุณลักษณะงานวิจัย และตรวจสอบคุณภาพ มีรายละเอียดดังนี้

2.1 ศึกษาตำรา เอกสาร และงานวิจัยที่เกี่ยวข้องกับการสร้างแบบบันทึกคุณลักษณะงานวิจัยเพื่อเป็นแนวทางในการกำหนดรูปแบบของแบบบันทึกและตัวแปรคุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

2.2 สร้างแบบบันทึกคุณลักษณะงานวิจัย ลักษณะเป็นแบบบันทึกข้อมูลที่มีการกำหนดรหัสสำหรับการบันทึกค่าในแต่ละรายการ พร้อมคู่มือลงรหัส

2.3 เมื่อสร้างเสร็จแล้วผู้วิจัยได้นำแบบบันทึกคุณลักษณะงานวิจัยไปเสนอต่อผู้เชี่ยวชาญจำนวน 5 ท่าน ซึ่งเป็นผู้เชี่ยวชาญกลุ่มเดียวกับที่พิจารณาแบบประเมินคุณภาพงานวิจัย เพื่อตรวจสอบความตรงเชิงเนื้อหา (content validity) และความครอบคลุมตามตัวแปรที่ต้องการจะบันทึก ผลการวิเคราะห์ความสอดคล้องและความเหมาะสมระหว่างคุณลักษณะงานวิจัยและรายละเอียดของคุณลักษณะงานวิจัย (IOC) มีค่าอยู่ระหว่าง 0.80-1.00 (รายละเอียดดังแสดงในภาคผนวก ข) อย่างไรก็ตามผู้เชี่ยวชาญได้ให้

ข้อเสนอแนะสำหรับปรับปรุงรายละเอียดของคุณลักษณะงานวิจัย ผู้วิจัยได้ดำเนินการปรับตามข้อเสนอแนะ มีรายละเอียดดังตารางที่ 3.5

ตารางที่ 3.5 ผลการปรับปรุงแบบบันทึกคุณลักษณะงานวิจัย และคู่มือลงรหัสตามข้อเสนอแนะของผู้เชี่ยวชาญ

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
<p>ควรแยก มหาวิทยาลัยของรัฐ กับ มหาวิทยาลัยในกำกับของรัฐ</p> <p>3. ประเภทสถาบันที่ผลิตงานวิจัย</p> <p>1 = มหาวิทยาลัยของรัฐ</p> <p>2 = มหาวิทยาลัยราชภัฏ</p> <p>3 = มหาวิทยาลัยเอกชน</p> <p>4 = สถาบันอื่น ๆ (รายงานวิจัย)</p>	<p>ผู้วิจัยปรับเกณฑ์การตามข้อเสนอแนะ ดังนี้</p> <p>3. ประเภทสถาบันที่ผลิตงานวิจัย</p> <p>1 = มหาวิทยาลัยของรัฐ</p> <p>2 = มหาวิทยาลัยในกำกับของรัฐ</p> <p>3 = มหาวิทยาลัยราชภัฏ</p> <p>4 = มหาวิทยาลัยเอกชน</p> <p>5 = สถาบันอื่น ๆ (รายงานวิจัย)</p>
<p>1) วัตถุประสงค์ข้อที่ 4 น่าจะอยู่ในข้อ 3</p> <p>2) วัตถุประสงค์ข้อที่ 5 และ 6 เหมือนกัน</p> <p>3) วัตถุประสงค์ข้อที่ 7 ยังไม่ชัด เป็น “เปรียบเทียบคุณลักษณะที่วัดกับตัวแปรอิสระ” หรือ “เปรียบเทียบคุณภาพ”</p> <p>8. วัตถุประสงค์ของการวิจัย</p> <p>(1) พัฒนาตัวบ่งชี้/องค์ประกอบ</p> <p>(2) สร้าง/พัฒนาเครื่องมือวัด</p> <p>(3) ตรวจสอบคุณภาพเครื่องมือวัด</p> <p>(4) ตรวจสอบความตรงเชิงโครงสร้าง</p> <p>(5) สร้างปกติวิสัย (norms)</p> <p>(6) พัฒนาเกณฑ์การประเมิน</p> <p>(7) เปรียบเทียบ</p> <p>(8) ทดสอบความไม่แปรเปลี่ยน</p>	<p>ผู้วิจัยปรับเกณฑ์การตามข้อเสนอแนะ ดังนี้</p> <p>8. วัตถุประสงค์ของการวิจัย</p> <p>(1) พัฒนาตัวบ่งชี้/องค์ประกอบ</p> <p>(2) สร้าง/พัฒนาเครื่องมือวัด</p> <p>(3) ตรวจสอบคุณภาพเครื่องมือวัด</p> <p>(4) สร้างปกติวิสัย (norms)</p> <p>(5) พัฒนาเกณฑ์การประเมิน/เกณฑ์การให้คะแนน</p> <p>(6) ศีรษะระดับ/เปรียบเทียบคุณลักษณะที่วัดกับตัวแปรอิสระ</p> <p>(7) ทดสอบความไม่แปรเปลี่ยน</p>
<p>การกำหนดนิยามเชิงปฏิบัติการ เพิ่มคำว่า “ตัวแปรหลักและองค์ประกอบที่วัด”</p> <p>20. การกำหนดนิยามเชิงปฏิบัติการ</p> <p>0 = ไม่มี</p> <p>1 = มี แต่ไม่นำเสนอ</p> <p>2 = มี และนำเสนอ</p>	<p>ผู้วิจัยปรับเกณฑ์การตามข้อเสนอแนะ ดังนี้</p> <p>20. การกำหนดนิยามเชิงปฏิบัติการของตัวแปรที่สำคัญและองค์ประกอบที่วัด</p> <p>0 = ไม่มี</p> <p>1 = มี แต่ไม่นำเสนอ</p> <p>2 = มี และนำเสนอ</p>
<p>เพิ่มข้อ แบบวัดเชิงสถานการณ์</p> <p>23. ประเภทของเครื่องมือวัด</p> <p>(1) แบบสอบหลายตัวเลือก</p> <p>(2) แบบสอบเติมคำ</p> <p>(3) แบบสอบความเรียง</p> <p>(4) แบบสอบภาคปฏิบัติ</p>	<p>ผู้วิจัยปรับเกณฑ์การตามข้อเสนอแนะ และเพิ่มเติมรายละเอียดให้สมบูรณ์ยิ่งขึ้น ดังนี้</p> <p>23. ประเภทของเครื่องมือวัด</p> <p>(1) แบบสอบหลายตัวเลือก</p> <p>(2) แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า</p> <p>(3) แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า</p>

ข้อเสนอแนะ	การปรับปรุงข้อเสนอแนะของผู้เชี่ยวชาญ
(5) แบบตรวจสอบรายการ	(4) แบบสอบอัตนัย
(6) มาตรฐานค่า	(5) แบบสอบภาคปฏิบัติ
(7) ประเภทอื่น ๆ	(6) แบบตรวจสอบรายการ
	(7) มาตรฐานค่า
	(8) แบบสังเกต
	(9) แบบสอบอัตนัยประยุกต์ (MEQ)
	(10) อื่น ๆ

2.4 คณะผู้วิจัยทดลองบันทึกคุณลักษณะงานวิจัยเพื่อตรวจสอบความชัดเจนของภาษาที่ใช้ รวมถึงความครอบคลุมของประเด็นที่ต้องการศึกษา แล้วนำมาปรับปรุงให้ครอบคลุมประเด็นที่ต้องการศึกษามากขึ้น

2.5 คณะผู้วิจัยทั้ง 2 คน นำแบบบันทึกคุณลักษณะงานวิจัยและคู่มือลงรหัสที่ปรับปรุงแล้ว มาร่วมบันทึกคุณลักษณะงานวิจัย จำนวน 3 เรื่อง ซึ่งเป็นเรื่องเดียวกันที่ใช้ในการประเมินคุณภาพงานวิจัย เพื่อตรวจสอบความสอดคล้องในการบันทึกคุณลักษณะงานวิจัย (inter rater reliability) จากการบันทึกคุณลักษณะงานวิจัยของผู้วิจัยทั้ง 2 คน นอกจากนี้ ผู้วิจัยยังได้ดำเนินการตรวจสอบความสอดคล้องภายในตัวผู้ประเมิน (intra-rater reliability) โดยผู้วิจัยลงบันทึกข้อมูลในแบบบันทึกคุณลักษณะงานวิจัย จากอ่านงานวิจัยทั้ง 3 เล่ม ซ้ำอีกครั้ง ทั้งระยะเวลาห่างกัน 2 สัปดาห์ และใช้สูตรของ Copper & Hedges (1994 อ้างถึงใน วิทยุ ยัมยวน, 2547) ในการพิจารณาความสอดคล้อง ผลการตรวจสอบความสอดคล้องในการบันทึกข้อมูลระหว่างผู้ประเมิน พบว่า มีค่าความสอดคล้องอยู่ระหว่าง 0.96-0.99 สรุปได้ว่า ผู้วิจัยทั้ง 2 คน สามารถบันทึกคุณลักษณะงานวิจัยได้สอดคล้องใกล้เคียงกันอยู่ในระดับดีมาก ในขณะที่ผลการตรวจสอบความสอดคล้องภายในตัวผู้ประเมิน พบว่า มีค่าความสอดคล้องอยู่ระหว่าง 0.99-1.00 จึงสรุปได้ว่าผู้วิจัยแต่ละคนสามารถบันทึกคุณลักษณะงานวิจัยได้ผลการบันทึกที่คงที่อยู่ในระดับดีมาก รายละเอียดดังตารางที่ 3.6

ตารางที่ 3.6 ผลการตรวจสอบความสอดคล้องระหว่างผู้ประเมินและภายในตัวผู้ประเมินจากการบันทึกคุณลักษณะงานวิจัย

เล่มงานวิจัย	จำนวนข้อที่มีการบันทึก*	ผู้วิจัยคนที่ 1 และ 2		ผู้วิจัยคนที่ 1 บันทึกผลครั้งที่ 2		ผู้วิจัยคนที่ 2 บันทึกผลครั้งที่ 2	
		จำนวนข้อที่บันทึกผลตรงกัน	AR	จำนวนข้อที่บันทึกผลตรงกัน	AR	จำนวนข้อที่บันทึกผลตรงกัน	AR
เล่มที่ 1	213	207	0.96	211	0.99	210	0.99
เล่มที่ 2	181	178	0.98	179	0.99	181	1.00
เล่มที่ 3	134	132	0.99	134	1.00	134	1.00

*หมายเหตุ: จำนวนข้อที่มีการบันทึกในแต่ละเล่มไม่เท่ากันเนื่องจากมีจำนวนเครื่องมือไม่เท่ากัน (AR: Agreement Rate, ค่าความสอดคล้องในการประเมิน)

การเก็บรวบรวมข้อมูล

ผู้วิจัยดำเนินการเก็บรวบรวมข้อมูล ดังนี้

1. สืบค้นงานวิจัย และคัดเลือกเฉพาะเล่มที่เป็นงานวิจัยฉบับสมบูรณ์ (full text) แต่ทั้งนี้ในกรณีที่ ไม่พบฉบับสมบูรณ์ ผู้วิจัยได้ดำเนินการสืบค้นไปยังฐานข้อมูลออนไลน์ของสถาบันนั้น ๆ ว่าสามารถสืบค้น ฉบับสมบูรณ์ได้หรือไม่

2. สืบค้นงานวิจัยในเบื้องต้น 1 รอบ เพื่อตรวจสอบเนื้อหาภายในตัวเล่มงานวิจัยนั้นว่า มีคุณสมบัติ เข้าเกณฑ์ในการคัดเลือกงานวิจัยที่ผู้วิจัยกำหนดไว้หรือไม่ และมีเนื้อหาสาระครบถ้วนเพียงพอที่จะนำมา สังเคราะห์งานวิจัยได้ เพื่อจะนำงานวิจัยฉบับนั้นไปใช้ในการประเมินคุณภาพงานวิจัย

3. อ่านงานวิจัยโดยละเอียด เพื่อประเมินคุณภาพงานวิจัย และบันทึกข้อมูลคุณลักษณะงานวิจัยลง ใบแบบบันทึก

4. จัดเตรียมข้อมูลสำหรับวิเคราะห์ ด้วยโปรแกรมสำเร็จรูป

5. สรุปองค์ความรู้ที่ได้จากการสังเคราะห์งานวิจัยด้วยการวิเคราะห์เนื้อหา

การวิเคราะห์ข้อมูลและสถิติที่ใช้

การวิเคราะห์ข้อมูลในครั้งนี้เป็นการสังเคราะห์งานวิจัยเชิงปริมาณด้วยการวิเคราะห์ทอิกิมาณ และการสังเคราะห์เชิงคุณภาพด้วยการวิเคราะห์เนื้อหา โดยแบ่งการวิเคราะห์เพื่อตอบวัตถุประสงค์การวิจัย ออกเป็น 4 ตอน คือ ตอนที่ 1 การประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนใน ศตวรรษที่ 21 ระดับมัธยมศึกษา ตอนที่ 2 การวิเคราะห์ข้อมูลเบื้องต้นงานวิจัยการพัฒนาเครื่องมือวัด สมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ตอนที่ 3 การวิเคราะห์คุณลักษณะงานวิจัยที่มีผล ต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา และ ตอนที่ 4 การสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษ ที่ 21 ระดับมัธยมศึกษา มีรายละเอียดดังนี้

ตอนที่ 1 การประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา เป็นการรายงานผลการประเมินคุณภาพงานวิจัย โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป ได้แก่ ค่าความถี่ ร้อยละ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน

ตอนที่ 2 การวิเคราะห์ข้อมูลเบื้องต้นงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนใน ศตวรรษที่ 21 ระดับมัธยมศึกษา เป็นการวิเคราะห์ข้อมูลเบื้องต้นโดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป ได้แก่ ค่าความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน สัมประสิทธิ์การกระจาย ค่าต่ำสุด ค่าสูงสุด ความเบ้ ความโด่งเพื่อให้ทราบลักษณะการแจกแจงการกระจายของตัวแปรคุณลักษณะงานวิจัย

ตอนที่ 3 การวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัด สมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

3.1 การวิเคราะห์ข้อมูลเบื้องต้นของดัชนีมาตรฐาน

3.2 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานจำแนกตามตัวแปร คุณลักษณะงานวิจัย โดยใช้การวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA)

ตอนที่ 4 การสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนใน ศตวรรษที่ 21 ระดับมัธยมศึกษา ด้วยการวิเคราะห์เนื้อหา

บทที่ 4 ผลการวิเคราะห์ข้อมูล

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อ 1) เพื่อประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา 2) เพื่อศึกษาคุณลักษณะของงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา และ 3) เพื่อสังเคราะห์องค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ทั้งนี้ ในการนำเสนอผลการวิเคราะห์ข้อมูล ผู้วิจัยแบ่งการนำเสนอเป็น 4 ตอน ได้แก่

ตอนที่ 1 ผลการประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

ตอนที่ 2 ผลการวิเคราะห์ข้อมูลเบื้องต้นงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษาที่นำมาสังเคราะห์

ตอนที่ 3 ผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

ตอนที่ 4 ผลสังเคราะห์องค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

มีรายละเอียดดังนี้

ตอนที่ 1 ผลการประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

ผลการวิเคราะห์ในตอนต้นที่ 1 เพื่อตอบวัตถุประสงค์ย่อยของการวิจัยข้อที่ 1 เพื่อประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ในตอนนี้ผู้วิจัยนำเสนอผลแบ่งออกเป็น 2 ส่วนคือ 1.1) ผลการประเมินระดับคุณภาพงานวิจัย และ 1.2) ผลการประเมินคุณภาพงานวิจัยรายด้าน มีรายละเอียดดังนี้

1.1 ผลการประเมินระดับคุณภาพงานวิจัย

งานวิจัยที่นำมาศึกษาในครั้งนี้มีจำนวน 92 เล่ม เป็นงานวิจัยในมหาวิทยาลัยในกำกับของรัฐ จำนวนมากที่สุด 43 เล่ม (ร้อยละ 46.74) รองลงมาคือ มหาวิทยาลัยราชภัฏ จำนวน 31 เล่ม (ร้อยละ 33.70) มหาวิทยาลัยของรัฐ จำนวน 11 เล่ม (ร้อยละ 11.96) และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 7 เล่ม (ร้อยละ 7.61) ตามลำดับ ผลการประเมินเพื่อจัดระดับคุณภาพ พบว่า ในภาพรวมส่วนใหญ่เป็นงานวิจัยที่มีคุณภาพอยู่ในระดับดี จำนวน 76 เล่ม คิดเป็นร้อยละ 82.61 ที่เหลือเป็นงานวิจัยที่มีคุณภาพอยู่ในระดับดีมาก จำนวน 16 เล่ม (ร้อยละ 17.39) ทั้งนี้ไม่มีงานวิจัยเล่มใดที่มีคุณภาพอยู่ในระดับปานกลาง คุณภาพค่อนข้างต่ำ และคุณภาพต่ำ

เมื่อพิจารณาจำแนกตามประเภทสถาบันที่ผลิตงานวิจัย พบว่า งานวิจัยที่มีคุณภาพในระดับดีมาก เป็นงานวิจัยจากมหาวิทยาลัยในกำกับของรัฐมากที่สุด จำนวน 12 เล่ม (ร้อยละ 13.04) รองลงมาคืองานวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 3 เล่ม (ร้อยละ 3.26) และมหาวิทยาลัยราชภัฏ จำนวน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ ในขณะที่งานวิจัยที่มีคุณภาพในระดับดี เป็นงานวิจัยจากมหาวิทยาลัยในกำกับของรัฐมากที่สุด จำนวน 31 เล่ม (ร้อยละ 33.70) รองลงมาคือ งานวิจัยจากมหาวิทยาลัยราชภัฏ จำนวน 30 เล่ม (ร้อยละ 32.61) มหาวิทยาลัยของรัฐ จำนวน 11 เล่ม (ร้อยละ

11.96) และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 4 เล่ม (ร้อยละ 4.35) ตามลำดับ

ทั้งนี้ เมื่อพิจารณาคุณภาพงานวิจัยตามตัวแปรประเภทงานวิจัย พบว่า งานวิจัยที่มีคุณภาพในระดับดีมาก เป็นงานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาโทมากที่สุด จำนวน 10 เล่ม (ร้อยละ 10.87) รองลงมาคือ งานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาเอก และรายงานวิจัย มีจำนวนเท่ากัน 3 เล่ม (ร้อยละ 3.26) ในขณะที่งานวิจัยที่มีคุณภาพในระดับดี เป็นงานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาโทมากที่สุด จำนวน 70 เล่ม (ร้อยละ 76.09) รองลงมาคือ รายงานวิจัย จำนวน 4 เล่ม (ร้อยละ 4.35) และงานวิจัยประเภทวิทยานิพนธ์ในระดับปริญญาเอก จำนวน 2 เล่ม (ร้อยละ 2.17) ตามลำดับ รายละเอียดดังตารางที่ 4.1

ตารางที่ 4.1 ระดับคุณภาพงานวิจัยจำแนกตามประเภทสถาบันที่ผลิตงานวิจัยและประเภทงานวิจัย

คุณลักษณะงานวิจัย	ดี		ดีมาก		รวม	
	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ
ประเภทสถาบันที่ผลิตงานวิจัย						
1. มหาวิทยาลัยของรัฐ	11	11.96	0	0.00	11	11.96
2. มหาวิทยาลัยในกำกับของรัฐ	31	33.70	12	13.04	43	46.74
3. มหาวิทยาลัยราชภัฏ	30	32.61	1	1.09	31	33.70
4. สทศ.	4	4.35	3	3.26	7	7.61
รวม	76	82.61	16	17.39	92	100.00
ประเภทงานวิจัย						
1. วิทยานิพนธ์ ป.โท	70	76.09	10	10.87	80	86.96
2. วิทยานิพนธ์ ป.เอก	2	2.17	3	3.26	5	5.43
รายงานวิจัย	4	4.35	3	3.26	7	7.61
รวม	76	82.61	16	17.39	92	100.00

1.2 ผลการประเมินคุณภาพงานวิจัยรายด้าน

เมื่อพิจารณาคุณภาพงานวิจัยในภาพรวม พบว่า มีคุณภาพอยู่ในระดับดี ($M=2.997$, $SD = 0.247$) เมื่อพิจารณาจำแนกรายด้าน พบว่า ด้านผลการวิเคราะห์ข้อมูล มีค่าเฉลี่ยสูงสุดมีคุณภาพอยู่ในระดับดีมาก ($M=3.875$, $SD = 0.382$) รองลงมาคือ ด้านการสรุปอภิปรายผลและข้อเสนอแนะ ($M=3.214$, $SD = 0.473$) ด้านวิธีดำเนินการวิจัย ($M= 3.046$, $SD = 0.370$) ด้านการนำเสนอรายงานวิจัย ($M= 2.783$, $SD = 0.340$) ด้านความเป็นมาและความสำคัญของปัญหาวิจัย ($M= 2.634$, $SD = 0.457$) และเอกสารและงานวิจัยที่เกี่ยวข้อง ($M= 2.429$, $SD = 0.340$) ตามลำดับ โดยมีคุณภาพอยู่ในระดับดีทุกด้าน ยกเว้น ด้านการสรุปอภิปรายผลและข้อเสนอแนะที่มีคุณภาพในระดับดีมาก

เมื่อพิจารณาคุณภาพงานวิจัยตามรายการประเมิน พบว่า คุณภาพงานวิจัยที่มีค่าสูงสุด 5 อันดับแรก มีคุณภาพอยู่ในระดับดีมากทุกรายการ ได้แก่ เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ ($M= 3.989$, $SD = 0.104$) รองลงมาคือ รูปแบบรายงานถูกต้องตามหลักวิชา ($M= 3.957$, $SD = 0.205$) ชื่อเรื่องมีความชัดเจน น่าสนใจ ($M= 3.946$, $SD = 0.228$) การแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องชัดเจน ($M= 3.946$, $SD = 0.310$) และลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูล ($M= 3.804$, $SD = 0.650$) ตามลำดับ ในขณะที่คุณภาพงานวิจัยในภาพรวมที่มีค่าต่ำสุด 5 อันดับแรก มีคุณภาพ

อยู่ในระดับค่อนข้างต่ำ 3 รายการ ได้แก่ เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย ($M = 1.130$, $SD = 0.667$) เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ ($M = 1.348$, $SD = 0.748$) และเหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล ($M = 1.565$, $SD = 0.816$) อีก 2 รายการมีคุณภาพในระดับปานกลาง ได้แก่ การนิยามศัพท์เฉพาะมีความชัดเจน ($M = 1.728$, $SD = 1.080$) และการสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา ($M = 2.196$, $SD = 1.361$) ตามลำดับ รายละเอียดดังตารางที่ 4.2

ตารางที่ 4.2 ผลการประเมินคุณภาพงานวิจัย

รายการประเมิน	M	SD	ระดับคุณภาพ
ความเป็นมาและความสำคัญของปัญหาวิจัย	2.634	0.457	ดี
1. ชื่อเรื่องมีความชัดเจน น่าสนใจ	3.946	0.228	ดีมาก
2. ความเป็นมาและความสำคัญของปัญหา ชัดเจน สอดคล้องกับเรื่องที่ทำวิจัย	3.022	0.851	ดี
3. ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง ถูกต้องตามหลักการวิจัย	3.163	0.700	ดี
4. เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล	1.565	0.816	ค่อนข้างต่ำ
5. กำหนดขอบเขตของการวิจัยกำหนดได้อย่างเหมาะสม และมีเหตุผลรองรับ	2.478	0.687	ดี
6. การนิยามศัพท์เฉพาะมีความชัดเจน	1.728	1.080	ปานกลาง
7. กรอบแนวคิดในการวิจัยมีความถูกต้องเหมาะสม ชัดเจน ตามหลักการวิจัย	2.533	1.190	ดี
เอกสารและงานวิจัยที่เกี่ยวข้อง	2.429	0.340	ดี
8. เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ	3.989	0.104	ดีมาก
9. เอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย	2.326	0.758	ปานกลาง
10. เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย	1.130	0.667	ค่อนข้างต่ำ
11. เอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศมีสัดส่วนที่เหมาะสม	2.272	0.537	ปานกลาง
วิธีดำเนินการวิจัย	3.046	0.370	ดี
12. การออกแบบการวิจัยสอดคล้องกับปัญหาการวิจัย	3.457	0.762	ดีมาก
13. ขั้นตอนการวิจัยมีความชัดเจน	3.054	0.906	ดี
14. กำหนดประชากรและตัวอย่างได้อย่างเหมาะสมกับการวิจัย	3.380	0.850	ดีมาก
15. ขนาดตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม	3.424	0.917	ดีมาก
16. การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน	2.239	1.133	ปานกลาง
17. การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีแนวคิดทฤษฎีเกี่ยวกับตัวแปรที่วัดและวิธีการสร้างเครื่องมือ	2.978	0.961	ดี
18. การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา	2.196	1.361	ปานกลาง
19. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ที่สร้างมีความเป็นปรนัยของข้อคำถาม	3.054	1.052	ดี

รายการประเมิน	M	SD	ระดับคุณภาพ
20. ขั้นตอนในการดำเนินการวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	3.054	0.942	ดี
21. วิธีการตรวจสอบคุณภาพเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	3.783	0.415	ดีมาก
22. มีการทดลองใช้เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 (try out)	2.870	1.112	ดี
23. ตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	3.707	0.688	ดีมาก
24. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ในภาพรวมมีคุณภาพเหมาะสม	2.793	0.989	ดี
25. กระบวนการเก็บรวบรวมข้อมูลมีความชัดเจนและเหมาะสม	3.043	0.982	ดี
26. สถิติที่ใช้ในการวิเคราะห์ข้อมูลมีความถูกต้องเหมาะสมกับลักษณะของข้อมูล	2.652	0.501	ดี
ผลการวิเคราะห์ข้อมูล	3.875	0.382	ดีมาก
27. ลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูล	3.804	0.650	ดีมาก
28. การแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องชัดเจน	3.946	0.310	ดีมาก
การสรุปอภิปรายผลและข้อเสนอแนะ	3.214	0.473	ดีมาก
29. สรุปผลการวิจัยอย่างถูกต้อง	3.467	0.523	ดีมาก
30. มีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีต	3.272	0.786	ดีมาก
31. ข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์	2.902	0.865	ดี
การนำเสนอรายงานวิจัย	2.783	0.340	ดี
32. เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางปฏิบัติ	3.043	0.417	ดี
33. เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ	1.348	0.748	ค่อนข้างต่ำ
34. รูปแบบรายงานถูกต้องตามหลักวิชา	3.957	0.205	ดีมาก
ภาพรวม	2.997	0.247	ดี

ตอนที่ 2 ผลการวิเคราะห์ข้อมูลเบื้องต้นงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษาที่นำมาสังเคราะห์

ผลการวิเคราะห์ข้อมูลเบื้องต้นงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษาที่นำมาสังเคราะห์ จำนวน 92 เล่ม แบ่งการนำเสนอออกเป็น 3 ส่วน คือ 2.1) ผลการวิเคราะห์ข้อมูลเบื้องต้นจำแนกตามตัวแปรคุณลักษณะงานวิจัยที่เป็นตัวแปรจัดประเภท 2.2) ผลการวิเคราะห์ข้อมูลเบื้องต้นจำแนกตามตัวแปรคุณลักษณะงานวิจัยที่เป็นตัวแปรต่อเนื่อง และ 2.3) ข้อมูลเบื้องต้นของการทดลองใช้เครื่องมือ มีรายละเอียดดังนี้

2.1 ผลการวิเคราะห์ข้อมูลเบื้องต้นจำแนกตามตัวแปรคุณลักษณะงานวิจัยที่เป็นตัวแปรจัดประเภท

ผลการวิเคราะห์ข้อมูลเบื้องต้นจำแนกตามตัวแปรคุณลักษณะงานวิจัยที่เป็นตัวแปรจัดประเภท แบ่งการนำเสนอตามตัวแปรคุณลักษณะของงานวิจัย 3 ด้าน ได้แก่ 2.1.1) ด้านการพิมพ์และผู้วิจัย, 2.1.2) ด้านเนื้อหาสาระงานวิจัย และ 2.1.3) ด้านคุณภาพงานวิจัย มีรายละเอียดดังนี้

2.1.1 ตัวแปรคุณลักษณะของงานวิจัยด้านการพิมพ์และผู้วิจัย

เมื่อพิจารณาตัวแปรคุณลักษณะของงานวิจัยด้านการพิมพ์และผู้วิจัย ประกอบด้วย ตัวแปรจัดประเภท 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, ประเภทงานวิจัย และเพศของผู้วิจัย จากงานวิจัยจำนวน 92 เล่ม ผลการวิเคราะห์ข้อมูลมีรายละเอียดดังนี้

ตัวแปรปีที่พิมพ์งานวิจัย พบว่า งานวิจัยมีการตีพิมพ์ในปี 2560 มากที่สุด จำนวน 10 เล่ม (ร้อยละ 10.87) รองลงมา 3 อันดับแรกคือ ปี 2555 และ 2558 ที่มีจำนวนเท่ากันคือ 9 เล่ม (ร้อยละ 9.78) และ ปี 2559 จำนวน 8 เล่ม (ร้อยละ 8.70) ตามลำดับ สำหรับปีที่พิมพ์งานวิจัยน้อยที่สุด คือ ปี 2546 จำนวน 1 เล่ม (ร้อยละ 1.09)

ตัวแปรประเภทสถาบันที่ผลิตงานวิจัย พบว่า เป็นงานวิจัยจากมหาวิทยาลัยในกำกับของรัฐ มากที่สุด จำนวน 43 เล่ม (ร้อยละ 46.74) รองลงมาคือ มหาวิทยาลัยราชภัฏ จำนวน 31 เล่ม (ร้อยละ 33.70), มหาวิทยาลัยของรัฐ จำนวน 11 เล่ม (ร้อยละ 11.96) และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 7 เล่ม (ร้อยละ 7.61) ตามลำดับ

ตัวแปรสถาบันที่ผลิตงานวิจัย พบว่า จุฬาลงกรณ์มหาวิทยาลัย และมหาวิทยาลัยราชภัฏมหาสารคาม เป็นสถาบันที่ผลิตงานวิจัยมากที่สุด จำนวนเท่ากัน คือ 14 เล่ม (ร้อยละ 15.22) รองลงมา 4 อันดับแรกคือ มหาวิทยาลัยเกษตรศาสตร์ และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ที่มีจำนวนเท่ากันคือ 7 เล่ม (ร้อยละ 7.61) มหาวิทยาลัยมหาสารคาม และมหาวิทยาลัยราชภัฏอุบลราชธานี มีจำนวนเท่ากันคือ 6 เล่ม (ร้อยละ 6.52) ตามลำดับ ในขณะที่สถาบันที่ผลิตงานวิจัยน้อยที่สุด คือ มหาวิทยาลัยราชภัฏเลย จำนวน 1 เล่ม (ร้อยละ 1.09)

ตัวแปรสาขาที่ผลิตงานวิจัย พบว่า เป็นกลุ่มสาขาการวัดและประเมินผล และกลุ่มสาขาการวิจัยและสถิติ มากที่สุด จำนวนเท่ากัน คือ 42 เล่ม (ร้อยละ 45.65) รองลงมา คือ รายงานวิจัย จำนวน 7 เล่ม (ร้อยละ 7.61) และสาขาเทคโนโลยีการวัดทางการศึกษา จำนวน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

ตัวแปรประเภทงานวิจัย พบว่า ส่วนใหญ่เป็นงานวิจัยระดับปริญญาโท จำนวน 80 เล่ม (ร้อยละ 86.96) รองลงมา เป็นรายงานวิจัย จำนวน 7 เล่ม (ร้อยละ 7.61) และงานวิจัยระดับปริญญาเอก จำนวน 5 เล่ม (ร้อยละ 5.43) ตามลำดับ

ตัวแปรเพศของผู้วิจัย พบว่า ส่วนใหญ่เป็นงานวิจัยที่ผลิตโดยนักวิจัยเพศหญิง จำนวน 76 เล่ม (ร้อยละ 82.61) รองลงมา คือ นักวิจัยเพศชาย จำนวน 15 เล่ม (ร้อยละ 16.30) และนักวิจัยชายและหญิง จำนวน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

2.1.2 ตัวแปรคุณลักษณะของงานวิจัยด้านเนื้อหาสาระงานวิจัย

เมื่อพิจารณาตัวแปรคุณลักษณะของงานวิจัยด้านเนื้อหาสาระงานวิจัย ประกอบด้วย ตัวแปรจัดประเภท 21 ตัวแปร ได้แก่ วัตถุประสงค์ของการวิจัย, เครื่องมือวัดที่พัฒนาขึ้น, แนวคิด/ทฤษฎีหลัก, องค์ประกอบที่มุ่งวัด, เนื้อหาที่เกี่ยวข้อง, ภูมิภาค, แหล่งที่มาของตัวอย่าง, ระดับชั้นของตัวอย่าง, การศึกษาแนวคิดการสร้างเครื่องมือวัด, การสร้างเครื่องมือวัด, การกำหนดนิยามเชิงปฏิบัติการ, การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด, การระบุขั้นตอนในการพัฒนาเครื่องมือวัด, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, จำนวนครั้งที่ใช้ในการทดลอง, ทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือ

วัด, การตรวจให้คะแนน, ประเภทของเกณฑ์ปกติ, วิธีที่ใช้ในการสร้างเกณฑ์ปกติ และคู่มือการใช้เครื่องมือวัด ผลการวิเคราะห์ข้อมูลมีรายละเอียดดังนี้

ตัวแปรวัตถุประสงค์ของการวิจัย พบว่า งานวิจัยมีวัตถุประสงค์เพื่อสร้าง/พัฒนาเครื่องมือวัด มากที่สุด จำนวน 89 เล่ม (ร้อยละ 33.46) รองลงมา 3 อันดับแรก คือ เพื่อตรวจสอบคุณภาพเครื่องมือวัด จำนวน 86 เล่ม (ร้อยละ 31.58), สร้างปกติวิสัย (norms) จำนวน 56 เล่ม (ร้อยละ 21.05) และ ศึกษาระดับ/เปรียบเทียบคุณลักษณะที่วัดกับตัวแปรอิสระ จำนวน 15 เล่ม (ร้อยละ 5.64) ตามลำดับ ในขณะที่วัตถุประสงค์ของการวิจัยมีการศึกษาน้อยที่สุด คือ พัฒนาเกณฑ์การประเมิน/เกณฑ์การให้คะแนน จำนวน 2 เล่ม (ร้อยละ 0.75)

ตัวแปรเครื่องมือวัดที่พัฒนาขึ้น พบว่า งานวิจัยมีการพัฒนาเครื่องมือวัดคุณลักษณะอันพึงประสงค์ มากที่สุด จำนวน 12 เล่ม (ร้อยละ 13.04) รองลงมา 4 อันดับแรก คือ เครื่องมือวัดความสามารถในการอ่าน คิดวิเคราะห์ และเขียนสื่อความ จำนวน 9 เล่ม (ร้อยละ 9.78), เครื่องมือวัดทักษะชีวิต/ทักษะชีวิตและอาชีพ, เครื่องมือวัดความสามารถการคิดอย่างมีวิจารณญาณ และเครื่องมือวัดทักษะการแก้ปัญหา ที่มีจำนวนเท่ากัน 8 เล่ม (ร้อยละ 8.70) ตามลำดับ ในขณะที่เครื่องมือวัดที่พัฒนาขึ้นน้อยที่สุด คือ เครื่องมืออื่น ๆ จำนวน 2 เล่ม (ร้อยละ 2.17)

ตัวแปรแนวคิด/ทฤษฎีหลัก พบว่า งานวิจัยใช้แนวคิด/ทฤษฎีหลักที่เกิดจากการสังเคราะห์ มากที่สุด จำนวน 29 เล่ม (ร้อยละ 31.52) รองลงมา 3 อันดับแรก คือ แนวคิด/ทฤษฎีหลักอื่น ๆ จำนวน 17 เล่ม (ร้อยละ 18.48) ม หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ.2551 จำนวน 14 เล่ม (ร้อยละ 15.22) และไม่ระบุ จำนวน 8 เล่ม (ร้อยละ 8.70) ตามลำดับ ในขณะที่แนวคิด/ทฤษฎีหลักที่มีการนำมาใช้น้อยที่สุด คือ แนวคิดการคิดอย่างมีวิจารณญาณของ Watson & Glaser, ทฤษฎีความคิดสร้างสรรค์ของ Torrance และทฤษฎีความคิดสร้างสรรค์ของ Guilford ที่จำนวนเท่ากัน 2 เล่ม (ร้อยละ 2.17)

ตัวแปรองค์ประกอบที่มุ่งวัด พบว่า งานวิจัยมีการศึกษาองค์ประกอบสมรรถนะของผู้เรียนในศตวรรษที่ 21 พบว่าเป็นการพัฒนาเครื่องมือวัดในองค์ประกอบ การสร้างสรรค์และนวัตกรรม มากที่สุด จำนวน 28 เล่ม (ร้อยละ 30.43) รองลงมา 3 อันดับแรก คือ การแก้ปัญหา จำนวน 25 เล่ม (ร้อยละ 27.17), การสื่อสาร จำนวน 23 เล่ม (ร้อยละ 25.00) และความรับผิดชอบ จำนวน 20 เล่ม (ร้อยละ 21.74) ตามลำดับ ในขณะที่องค์ประกอบที่มีการพัฒนาเครื่องมือวัดน้อยที่สุด คือ ความรู้เรื่องพลเมือง จำนวน 1 เล่ม (ร้อยละ 1.09)

ตัวแปรเนื้อหาที่เกี่ยวข้อง พบว่า งานวิจัยมีการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ที่ไม่อิงเนื้อหาวิชา (content free) มากที่สุด จำนวน 56 เล่ม (ร้อยละ 60.87) รองลงมา 3 อันดับแรก คือ กลุ่มสาระวิทยาศาสตร์ จำนวน 14 เล่ม (ร้อยละ 15.22), คณิตศาสตร์ จำนวน 13 เล่ม (ร้อยละ 14.13) และภาษาต่างประเทศ จำนวน 5 เล่ม (ร้อยละ 5.43) ตามลำดับ ในขณะที่มีการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 น้อยที่สุด ในกลุ่มสังคมศึกษา ศาสนา และวัฒนธรรม, ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์), การงานอาชีพและเทคโนโลยี และมากกว่า 1 สาระ ที่จำนวนเท่ากัน 1 เล่ม (ร้อยละ 1.09)

ตัวแปรภูมิภาค พบว่า งานวิจัยมีการศึกษาในภูมิภาคตะวันออกเฉียงเหนือมากที่สุด จำนวน 35 เล่ม (ร้อยละ 38.04) รองลงมาคือ ภาคกลาง จำนวน 18 เล่ม (ร้อยละ 19.57), ภาคเหนือ จำนวน 13 เล่ม (ร้อยละ 14.13), กรุงเทพมหานคร จำนวน 12 เล่ม (ร้อยละ 13.04), ภาคใต้ จำนวน 7 เล่ม

(ร้อยละ 7.61), มากกว่า 1 ภูมิภาค จำนวน 6 เล่ม (ร้อยละ 6.52) และไม่ระบุ จำนวน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

ตัวแปรแหล่งที่มาของตัวอย่าง พบว่า งานวิจัยศึกษากับกลุ่มตัวอย่างจากสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) มากที่สุด จำนวน 76 เล่ม (ร้อยละ 82.61) รองลงมา คือ ไม่ระบุ จำนวน 7 เล่ม (ร้อยละ 7.61), มากกว่า 1 สังกัด จำนวน 3 เล่ม (ร้อยละ 3.26), สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน (สช.) และองค์กรปกครองส่วนท้องถิ่น (อปท.) มีจำนวนเท่ากัน 2 เล่ม (ร้อยละ 2.17) สังกัดคณะกรรมการการอุดมศึกษา (สกอ.) และอื่น ๆ มีจำนวนเท่ากัน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

ตัวแปรระดับชั้นของตัวอย่าง พบว่า งานวิจัยศึกษากับนักเรียนชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน) มากที่สุด จำนวน 23 เล่ม (ร้อยละ 25.00) รองลงมา 3 อันดับแรก คือ ชั้นมัธยมศึกษาปีที่ 3 จำนวน 19 เล่ม (ร้อยละ 20.65), ชั้นมัธยมศึกษาปีที่ 2 และชั้นมัธยมศึกษาตอนปลาย (มากกว่า 1 ชั้นเรียน) มีจำนวนเท่ากัน 10 เล่ม (ร้อยละ 10.87) ตามลำดับ ในขณะที่มีการศึกษากับนักเรียนชั้นประถมศึกษาและมัธยมศึกษา และชั้นมัธยมศึกษาปีที่ 4 น้อยที่สุด มีจำนวนเท่ากัน 2 เล่ม (ร้อยละ 2.17)

ตัวแปรการศึกษาแนวความคิดการสร้างเครื่องมือวัด พบว่า งานวิจัยมีการศึกษาทั้งวิธีการสร้างและการวิเคราะห์คุณภาพ มากที่สุด จำนวน 65 เล่ม (ร้อยละ 70.65) รองลงมา มีการศึกษาเฉพาะวิธีการสร้างเครื่องมือ จำนวน 18 เล่ม (ร้อยละ 19.57) และมีการศึกษาเฉพาะการวิเคราะห์คุณภาพเครื่องมือ จำนวน 9 เล่ม (ร้อยละ 9.78) ตามลำดับ

ตัวแปรการสร้างเครื่องมือวัด พบว่า งานวิจัยเกือบทั้งหมดผู้วิจัยเป็นผู้พัฒนาเครื่องมือวัดขึ้นเอง จำนวน 91 เล่ม (ร้อยละ 98.91) และมีเพียง 1 เล่ม (ร้อยละ 1.09) ที่ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยต่างประเทศ

ตัวแปรการกำหนดนิยามเชิงปฏิบัติการ พบว่า งานวิจัยส่วนใหญ่มีและนำเสนอนิยามเชิงปฏิบัติการ จำนวน 76 เล่ม (ร้อยละ 82.61) รองลงมา มีการนิยามเชิงปฏิบัติการแต่ไม่นำเสนอ จำนวน 10 เล่ม (ร้อยละ 10.87) และไม่มีการกำหนดนิยามเชิงปฏิบัติการ จำนวน 6 เล่ม (ร้อยละ 6.52) ตามลำดับ

ตัวแปรการสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด พบว่า งานวิจัยส่วนใหญ่มีการสร้างตารางกำหนดโครงสร้างของเครื่องมือวัดและนำเสนอไว้ในงานวิจัย จำนวน 65 เล่ม (ร้อยละ 70.65) รองลงมาคือ ไม่มีการสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด จำนวน 17 เล่ม (ร้อยละ 18.48) และมีการสร้างตารางกำหนดโครงสร้างของเครื่องมือวัดแต่ไม่ได้นำเสนอไว้ในงานวิจัย จำนวน 10 เล่ม (ร้อยละ 10.87) ตามลำดับ

ตัวแปรการระบุขั้นตอนในการพัฒนาเครื่องมือวัด พบว่า งานวิจัยทุกเล่มมีการระบุขั้นตอนในการพัฒนาเครื่องมือวัด (ร้อยละ 100)

ตัวแปรประเภทของเครื่องมือวัด พบว่า งานวิจัยมีการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 เป็นแบบสอบถามอัตนัย มากที่สุด จำนวน 27 เล่ม (ร้อยละ 29.35) รองลงมา 3 อันดับแรก คือ มาตราประมาณค่า จำนวน 25 เล่ม (ร้อยละ 27.17), แบบวัดเชิงสถานการณ์ แบบให้คะแนน มากกว่า 2 ค่า จำนวน 23 เล่ม (ร้อยละ 25.00) และแบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า จำนวน 19 เล่ม (ร้อยละ 20.65) ตามลำดับ ในขณะที่มีการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 เป็นแบบตรวจสอบรายการ และอื่นๆ น้อยที่สุด จำนวนเท่ากัน 1 เล่ม (ร้อยละ 1.09)

ตัวแปรจำนวนตัวเลือก/เกณฑ์ พบว่า เครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 เป็นเครื่องมือที่มี 4 ตัวเลือก/เกณฑ์ มากที่สุด จำนวน 28 เล่ม (ร้อยละ 30.43) รองลงมา คือ เป็น

เครื่องมือที่มีตัวเลือก/เกณฑ์มากกว่า 1 ประเภท จำนวน 25 เล่ม (ร้อยละ 27.17), ไม่มีตัวเลือก/เกณฑ์ จำนวน 17 เล่ม (ร้อยละ 18.48), 5 ตัวเลือก/เกณฑ์ จำนวน 14 เล่ม (ร้อยละ 15.22), 3 ตัวเลือก/เกณฑ์ จำนวน 6 เล่ม (ร้อยละ 6.52) และมากกว่า 5 ตัวเลือก/เกณฑ์ จำนวน 2 เล่ม (ร้อยละ 2.17) ตามลำดับ

ตัวแปรจำนวนครั้งที่ใช้ในการทดลอง พบว่า งานวิจัยมีการทดลองใช้เครื่องมือจำนวน 3 ครั้ง มากที่สุด จำนวน 41 เล่ม (ร้อยละ 44.57) รองลงมา คือ จำนวน 2 ครั้ง จำนวน 40 เล่ม (ร้อยละ 43.48), จำนวน 4 ครั้ง จำนวน 7 เล่ม (ร้อยละ 7.61) และจำนวน 1 ครั้ง จำนวน 2 เล่ม (ร้อยละ 2.17) ตามลำดับ

ตัวแปรทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือวัด พบว่า เป็นงานวิจัยที่ใช้ทฤษฎีการทดสอบแบบดั้งเดิม (CTT) มากที่สุด จำนวน 78 เล่ม (ร้อยละ 84.78) รองลงมา คือ ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีตอบสนองข้อสอบ (CCT & IRT) จำนวน 8 เล่ม (ร้อยละ 8.70), ทฤษฎีการทดสอบแบบดั้งเดิมและทฤษฎีการสรุปอ้างอิงการทดสอบ (CCT & G Theory) (CCT & IRT) จำนวน 4 เล่ม (ร้อยละ 4.35), ทฤษฎีตอบสนองข้อสอบ (IRT) และทฤษฎีการสรุปอ้างอิงการทดสอบและทฤษฎีตอบสนองข้อสอบ (G Theory & IRT) มีจำนวนเท่ากัน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

ตัวแปรการตรวจให้คะแนน พบว่า เครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 มีการให้คะแนนมากกว่าสองค่า มากที่สุด จำนวน 52 เล่ม (ร้อยละ 56.52) รองลงมา คือ การให้คะแนนแบบสองค่า และมากกว่าสองค่า จำนวน 23 เล่ม (ร้อยละ 25.00) และให้คะแนนแบบสองค่า จำนวน 17 เล่ม (ร้อยละ 18.48) ตามลำดับ

ตัวแปรประเภทของเกณฑ์ปกติ พบว่า งานวิจัยมีการสร้างเกณฑ์ปกติระดับท้องถิ่นมากที่สุด จำนวน 33 เล่ม (ร้อยละ 35.87) รองลงมา คือ ไม่มีการสร้างเกณฑ์ปกติ จำนวน 30 เล่ม (ร้อยละ 36.21), ไม่ระบุ จำนวน 25 เล่ม (ร้อยละ 27.17) นอกนั้นเป็นเกณฑ์ปกติภายในชั้นเรียน เกณฑ์ปกติระดับสถาบัน เกณฑ์ปกติระดับชาติ และเกณฑ์ปกติระดับท้องถิ่น และเกณฑ์ปกติระดับชาติ มีจำนวนเท่ากัน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

ตัวแปรวิธีที่ใช้ในการสร้างเกณฑ์ปกติ พบว่า งานวิจัยใช้วิธีที่ใช้ในการสร้างเกณฑ์ปกติด้วยวิธีคะแนนมาตรฐานที่ปกติ มากที่สุด จำนวน 54 เล่ม (ร้อยละ 58.70) รองลงมาคือ ไม่มีการสร้างเกณฑ์ปกติ จำนวน 30 เล่ม (ร้อยละ 36.21), วิธีเปอร์เซ็นต์ไทล์ และวิธีเปอร์เซ็นต์ไทล์และสเตโนไนน์รวมกัน มีจำนวนเท่ากัน 3 เล่ม (ร้อยละ 3.26) นอกนั้นเป็นวิธีคะแนนมาตรฐานที่ และวิธีเปอร์เซ็นต์ไทล์และคะแนนมาตรฐานที่รวมกัน มีจำนวนเท่ากัน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

ตัวแปรคู่มือการใช้เครื่องมือวัด พบว่า งานวิจัยมีการสร้างคู่มือการใช้เครื่องมือวัด จำนวน 70 เล่ม (ร้อยละ 76.09) และไม่มีการสร้างคู่มือการใช้เครื่องมือวัด จำนวน 22 เล่ม (ร้อยละ 23.91)

2.1.3 ตัวแปรคุณลักษณะของงานวิจัยด้านวิธีวิทยาการวิจัย

เมื่อพิจารณาตัวแปรคุณลักษณะของงานวิจัยด้านวิธีวิทยาการวิจัย ประกอบด้วย ตัวแปรจัดประเภท 3 ตัวแปร ได้แก่ การกำหนดขนาดของตัวอย่าง, วิธีการเลือกตัวอย่าง และประเภทสถิติวิเคราะห์ข้อมูล ผลการวิเคราะห์ข้อมูลมีรายละเอียดดังนี้

ตัวแปรการกำหนดขนาดของตัวอย่าง พบว่า งานวิจัยใช้สูตรของ Yamane (1973) มากที่สุด จำนวน 20 เล่ม (ร้อยละ 21.74) รองลงมา 3 อันดับแรก คือ 20 เท่าของตัวแปร/พารามิเตอร์ จำนวน 14 เล่ม (ร้อยละ 15.22), หลาย ๆ เกณฑ์ จำนวน 9 เล่ม (ร้อยละ 9.78) และสูตรของ Krejcie &

Morgan (1970) จำนวน 8 เล่ม (ร้อยละ 8.70) ในขณะที่วิธีการกำหนดขนาดของตัวอย่างที่งานวิจัยใช้น้อยที่สุด คือ 15 เท่าของตัวแปร/พารามิเตอร์ จำนวน 1 เล่ม (ร้อยละ 1.09)

ตัวแปรวิธีการเลือกตัวอย่าง พบว่า งานวิจัยใช้สุ่มแบบหลายขั้นตอน มากที่สุด จำนวน 46 เล่ม (ร้อยละ 50.00) รองลงมาคือ สุ่มแบบสองขั้นตอน จำนวน 20 เล่ม (ร้อยละ 21.74), สุ่มแบบแบ่งชั้น จำนวน 12 เล่ม (ร้อยละ 13.04), ไม่ระบุ จำนวน 6 เล่ม (ร้อยละ 6.52), เลือกอย่างเจาะจง จำนวน 5 เล่ม (ร้อยละ 5.43), สุ่มอย่างง่าย จำนวน 2 เล่ม (ร้อยละ 2.17) และสุ่มแบบกลุ่ม จำนวน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

ตัวแปรประเภทสถิติวิเคราะห์ข้อมูล พบว่า งานวิจัยมีการวิเคราะห์ค่าความเที่ยง และค่าความตรงทุกเล่ม (ร้อยละ 100) รองลงมา 3 อันดับแรก คือ ค่าอำนาจจำแนก จำนวน 89 เล่ม (ร้อยละ 96.74), descriptive stat จำนวน 67 เล่ม (ร้อยละ 72.83) และค่าความยาก จำนวน 51 เล่ม (ร้อยละ 55.43) ในขณะที่ MANOVA น้อยที่สุด จำนวนเท่ากัน 1 เล่ม (ร้อยละ 1.09) ตามลำดับ

2.1.4 ตัวแปรคุณลักษณะของงานวิจัยด้านคุณภาพงานวิจัย

เมื่อพิจารณาตัวแปรคุณลักษณะของงานวิจัยด้านคุณภาพงานวิจัย มีตัวแปรจัดประเภท 1 ตัวแปร คือ ระดับคุณภาพงานวิจัย ผลการวิเคราะห์ข้อมูล พบว่า งานวิจัยส่วนใหญ่มีคุณภาพอยู่ในระดับดี จำนวน 76 เล่ม (ร้อยละ 82.61) และคุณภาพในระดับดีมาก 16 เล่ม (ร้อยละ 17.39) รายละเอียดดังตารางที่ 4.3

ตารางที่ 4.3 ข้อมูลพื้นฐานคุณลักษณะงานวิจัยระดับเล่มงานวิจัยที่เป็นตัวแปรจัดประเภท

ตัวแปรคุณลักษณะของงานวิจัย	ความถี่	ร้อยละ
ด้านการพิมพ์และผู้วิจัย		
ปีที่พิมพ์งานวิจัย		
(1) 2546	1	1.09
(2) 2547	4	4.35
(3) 2548	7	7.61
(4) 2549	6	6.52
(5) 2550	5	5.43
(6) 2551	4	4.35
(7) 2552	5	5.43
(8) 2553	6	6.52
(9) 2554	3	3.26
(10) 2555	9	9.78
(11) 2556	7	7.61
(12) 2557	5	5.43
(13) 2558	9	9.78
(14) 2559	8	8.70
(15) 2560	10	10.87
(16) 2561	3	3.26
รวม	92	100
ประเภทสถาบันที่ผลิตงานวิจัย		
(1) มหาวิทยาลัยของรัฐ	11	11.96

ตัวแปรคุณลักษณะของงานวิจัย	ความถี่	ร้อยละ
(2) มหาวิทยาลัยในกำกับของรัฐ	43	46.74
(3) มหาวิทยาลัยราชภัฏ	31	33.70
(4) สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)	7	7.61
รวม	92	100
สถาบันที่ผลิตงานวิจัย		
(1) จุฬาลงกรณ์มหาวิทยาลัย	14	15.22
(2) มหาวิทยาลัยเกษตรศาสตร์	7	7.61
(3) มหาวิทยาลัยเชียงใหม่	3	3.26
(4) มหาวิทยาลัยขอนแก่น	4	4.35
(5) มหาวิทยาลัยทักษิณ	5	5.43
(6) มหาวิทยาลัยนเรศวร	2	2.17
(7) มหาวิทยาลัยบูรพา	4	4.35
(8) มหาวิทยาลัยมหาสารคาม	6	6.52
(9) มหาวิทยาลัยศรีนครินทรวิโรฒ	4	4.35
(10) มหาวิทยาลัยสงขลานครินทร์	2	2.17
(11) มหาวิทยาลัยสุโขทัยธรรมาธิราช	4	4.35
(12) มหาวิทยาลัยราชภัฏกาญจนบุรี	3	3.26
(13) มหาวิทยาลัยราชภัฏนครราชสีมา	3	3.26
(14) มหาวิทยาลัยราชภัฏเพชรบูรณ์	3	3.26
(15) มหาวิทยาลัยราชภัฏมหาสารคาม	14	15.22
(16) มหาวิทยาลัยราชภัฏเลย	1	1.09
(17) มหาวิทยาลัยราชภัฏอุบลราชธานี	6	6.52
(18) สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน)	7	7.61
รวม	92	100
สาขาที่ผลิตงานวิจัย		
(1) กลุ่มสาขาการวัดและประเมินผล (การวัดผลการศึกษา, การวัดและประเมินผลการศึกษา)	42	45.65
(2) กลุ่มสาขาการวิจัยและสถิติ (วิจัยและประเมินผลการศึกษา, การวัดผลและการวิจัยการศึกษา, วัดผลและสถิติการศึกษา)	42	45.65
(3) เทคโนโลยีการวัดทางการศึกษา	1	1.09
(4) รายงานวิจัย	7	7.61
รวม	92	100
ประเภทงานวิจัย		
(1) วิทยานิพนธ์ ปริญญาโท	80	86.96
(2) วิทยานิพนธ์ ปริญญาเอก	5	5.43
(3) รายงานวิจัย	7	7.61
รวม	92	100
เพศของผู้วิจัย		
(1) ชาย	15	16.30
(2) หญิง	76	82.61
(3) ชายและหญิง	1	1.09
รวม	92	100

ตัวแปรคุณลักษณะของงานวิจัย	ความถี่	ร้อยละ
ด้านเนื้อหาสาระงานวิจัย		
วัตถุประสงค์ของการวิจัย*		
(1) พัฒนาตัวบ่งชี้/องค์ประกอบ	5	1.88
(2) สร้าง/พัฒนาเครื่องมือวัด	89	33.46
(3) ตรวจสอบคุณภาพเครื่องมือวัด	84	31.58
(4) สร้างปกติวิสัย (norms)	56	21.05
(5) พัฒนาเกณฑ์การประเมิน/เกณฑ์การให้คะแนน	2	0.75
(6) ศึกษาระดับ/เปรียบเทียบคุณลักษณะที่วัดกับตัวแปรอิสระ	15	5.64
(7) ทดสอบความไม่แปรเปลี่ยน	3	1.13
(8) สร้างคู่มือการใช้เครื่องมือวัด	12	4.51
รวม	266	100
เครื่องมือวัดที่พัฒนาขึ้น		
(1) เครื่องมือวัดทักษะแห่งศตวรรษที่ 21	4	4.35
(2) เครื่องมือวัดความตระหนักรู้ต่อโลก	3	3.26
(3) เครื่องมือวัดคุณลักษณะอันพึงประสงค์	12	13.04
(4) เครื่องมือวัดความมีวินัยในตนเอง	3	3.26
(5) เครื่องมือวัดทักษะชีวิต/ทักษะชีวิตและอาชีพ	8	8.70
(6) เครื่องมือวัดทักษะการคิด/ทักษะการคิดขั้นสูง	8	8.70
(7) เครื่องมือวัดความสามารถการคิดอย่างมีวิจารณญาณ	8	8.70
(8) เครื่องมือวัดทักษะการแก้ปัญหา	8	8.70
(9) เครื่องมือวัดความคิดสร้างสรรค์	6	6.52
(10) เครื่องมือวัดทักษะกระบวนการทางคณิตศาสตร์	5	5.43
(11) เครื่องมือวัดจิตวิทยาศาสตร์	7	7.61
(12) เครื่องมือวัดความสามารถในการอ่าน คิดวิเคราะห์ และเขียนสื่อความ	9	9.78
(13) เครื่องมือวัดความสามารถทางด้านเทคโนโลยีสารสนเทศ	5	5.43
(14) เครื่องมือวัดทักษะการรู้เท่าทันสื่อ	4	4.35
(15) อื่น ๆ	2	2.17
รวม	92	100
แนวคิด/ทฤษฎีหลัก		
(1) Partnership for 21st Century Learning	4	4.35
(2) หลักสูตรการศึกษาขั้นพื้นฐาน พ.ศ.2544	4	4.35
(3) หลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พ.ศ.2551	14	15.22
(4) แนวคิดทักษะชีวิตของกรมสุขภาพจิต	4	4.35
(5) แนวคิดการคิดอย่างมีวิจารณญาณของ Watson & Glaser	2	2.17
(6) แนวคิดการคิดอย่างมีวิจารณญาณของ Ennis (1985)	4	4.35
(7) ทฤษฎีความคิดสร้างสรรค์ของ Torrance	2	2.17
(8) ทฤษฎีความคิดสร้างสรรค์ของ Guilford	2	2.17
(9) คุณลักษณะจิตวิทยาศาสตร์ ของ สสวท.	2	2.17
(10) สังเคราะห์จากหลายแนวคิด/ทฤษฎี	29	31.52
(11) อื่น ๆ	17	18.48
(12) ไม่ระบุ	8	8.70
รวม	92	100

	ตัวแปรคุณลักษณะของงานวิจัย	ความถี่	ร้อยละ
องค์ประกอบที่มุ่งวัด			
(1) จิตสำนึกต่อโลก	ใช่	3	3.30
	ไม่ใช่	89	96.70
	รวม	92	100
(2) ความรู้เรื่องพลเมือง	ใช่	1	1.09
	ไม่ใช่	91	98.91
	รวม	92	100
(3) การคิดอย่างมีวิจารณญาณ	ใช่	16	17.39
	ไม่ใช่	76	82.61
	รวม	92	100
(4) การแก้ปัญหา	ใช่	25	27.17
	ไม่ใช่	67	72.83
	รวม	92	100
(5) การสร้างสรรค์และนวัตกรรม	ใช่	28	30.43
	ไม่ใช่	64	69.57
	รวม	92	100
(6) การสื่อสาร	ใช่	23	25.00
	ไม่ใช่	69	75.00
	รวม	92	100
(7) ทักษะด้านการทำงานเป็นทีม ความร่วมมือ	ใช่	8	8.70
	ไม่ใช่	84	91.30
	รวม	92	100
(8) ภาวะผู้นำ	ใช่	4	4.35
	ไม่ใช่	88	95.65
	รวม	92	100
(9) ความรับผิดชอบ	ใช่	20	21.74
	ไม่ใช่	72	78.26
	รวม	92	100
(10) การรู้เท่าทันสื่อและสารสนเทศ	ใช่	6	6.52
	ไม่ใช่	86	93.48
	รวม	92	100
(11) การรู้เท่าทันเทคโนโลยี	ใช่	8	8.70
	ไม่ใช่	84	91.30
	รวม	92	100
(12) การคิดอย่างมีวิจารณญาณและการแก้ปัญหา	ใช่	4	4.35
	ไม่ใช่	88	95.65
	รวม	92	100
(13) การแก้ปัญหอย่างสร้างสรรค์	ใช่	2	2.17
	ไม่ใช่	90	97.83
	รวม	92	100

ตัวแปรคุณลักษณะของงานวิจัย		ความถี่	ร้อยละ
(14) การสื่อสารและการร่วมมือ	ใช่	3	3.26
	ไม่ใช่	89	96.74
	รวม	92	100
(15) ภาวะผู้นำและความรับผิดชอบ	ใช่	2	2.17
	ไม่ใช่	90	97.83
	รวม	92	100
เนื้อหาที่เกี่ยวข้อง			
(1) คณิตศาสตร์		13	14.13
(2) วิทยาศาสตร์		14	15.22
(3) สังคมศึกษา ศาสนา และวัฒนธรรม		1	1.09
(4) ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์)		1	1.09
(5) การงานอาชีพและเทคโนโลยี		1	1.09
(6) ภาษาต่างประเทศ		5	5.43
(7) ไม่มีเนื้อหาวิชา (content free)		56	60.87
(8) มากกว่า 1 สาระ		1	1.09
รวม		92	100
ภูมิภาค			
(1) กทม.		12	13.04
(2) ภาคเหนือ		13	14.13
(3) ภาคตะวันออกเฉียงเหนือ		35	38.04
(4) ภาคกลาง		18	19.57
(5) ภาคใต้		7	7.61
(6) มากกว่า 1 ภูมิภาค		6	6.52
(7) ไม่ระบุ		1	1.09
รวม		92	100
แหล่งที่มาของตัวอย่าง			
(1) สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.)		76	82.61
(2) สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน		2	2.17
(3) สังกัดองค์กรปกครองส่วนท้องถิ่น		2	2.17
(4) สังกัดคณะกรรมการการอุดมศึกษา (สกอ.)		1	1.09
(5) อื่น ๆ		1	1.09
(6) มากกว่า 1 สังกัด		3	3.26
(7) ไม่ระบุ		7	7.61
รวม		92	100
ระดับชั้นของตัวอย่าง			
(1) ชั้นมัธยมศึกษาปีที่ 1		8	8.70
(2) ชั้นมัธยมศึกษาปีที่ 2		10	10.87
(3) ชั้นมัธยมศึกษาปีที่ 3		19	20.65
(4) ชั้นมัธยมศึกษาปีที่ 4		2	2.17
(5) ชั้นมัธยมศึกษาปีที่ 5		8	8.70
(6) ชั้นมัธยมศึกษาปีที่ 6		3	3.26
(7) ชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน)		23	25.00

ตัวแปรคุณลักษณะของงานวิจัย	ความถี่	ร้อยละ	
(8) ชั้นมัธยมศึกษาตอนปลาย (มากกว่า 1 ชั้นเรียน)	10	10.87	
(9) ชั้นมัธยมศึกษาตอนต้นและตอนปลาย	7	7.61	
(10) ชั้นประถมศึกษาและมัธยมศึกษา	2	2.17	
รวม	92	100	
การศึกษาแนวคิดการสร้างเครื่องมือวัด			
(1) มี เฉพาะการวิเคราะห์คุณภาพเครื่องมือ	9	9.78	
(2) มี เฉพาะวิธีการสร้างเครื่องมือ	18	19.57	
(3) มี ทั้งวิธีการสร้างและการวิเคราะห์คุณภาพ	65	70.65	
รวม	92	100	
การสร้างเครื่องมือวัด			
(1) ผู้วิจัยพัฒนาเครื่องมือวัดขึ้นเอง	91	98.91	
(2) ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยต่างประเทศ	1	1.09	
รวม	92	100	
การกำหนดนิยามเชิงปฏิบัติการ			
(1) ไม่มี	6	6.52	
(2) มี แต่ไม่นำเสนอ	10	10.87	
(3) มี และนำเสนอ	76	82.61	
รวม	92	100	
การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด			
(1) ไม่มี	17	18.48	
(2) มี แต่ไม่นำเสนอ	10	10.87	
(3) มี และนำเสนอ	65	70.65	
รวม	92	100	
การระบุขั้นตอนในการพัฒนาเครื่องมือวัด			
มี	92	100	
ไม่มี	0	0	
รวม	92	100	
ประเภทของเครื่องมือวัด			
(1) แบบสอบถามหลายตัวเลือก	ใช่	17	18.48
	ไม่ใช่	75	81.52
	รวม	92	100
(2) แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า	ใช่	19	20.65
	ไม่ใช่	73	79.35
	รวม	92	100
(3) แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า	ใช่	23	25.00
	ไม่ใช่	69	75.00
	รวม	92	100
(4) แบบสอบอัตนัย	ใช่	27	29.35
	ไม่ใช่	65	70.65
	รวม	92	100
(5) แบบสอบภาคปฏิบัติ	ใช่	6	6.52
	ไม่ใช่	86	93.48
	รวม	92	100

ตัวแปรคุณลักษณะของงานวิจัย		ความถี่	ร้อยละ
(6) แบบตรวจสอบรายการ	ใช่	1	1.09
	ไม่ใช่	91	98.91
	รวม	92	100
(7) มาตรฐานค่า	ใช่	25	27.17
	ไม่ใช่	67	72.83
	รวม	92	100
(8) แบบสังเกต	ใช่	3	3.26
	ไม่ใช่	89	96.74
	รวม	92	100
(9) แบบสอบถามนัยประยุกต์ (MEQ)	ใช่	5	5.43
	ไม่ใช่	87	94.57
	รวม	92	100
(10) อื่น ๆ	ใช่	1	1.09
	ไม่ใช่	91	98.9
	รวม	92	100
จำนวนตัวเลือก/เกณฑ์			
(1) ไม่มีตัวเลือก/เกณฑ์		17	18.48
(2) 3 ตัวเลือก/เกณฑ์		6	6.52
(3) 4 ตัวเลือก/เกณฑ์		28	30.43
(4) 5 ตัวเลือก/เกณฑ์		14	15.22
(5) มากกว่า 5 ตัวเลือก/เกณฑ์		2	2.17
(6) ใช้มากกว่า 1 ประเภท		25	27.17
	รวม	92	100
จำนวนครั้งที่ใช้ในการทดลอง			
(1) 1 ครั้ง		2	2.17
(2) 2 ครั้ง		40	43.48
(3) 3 ครั้ง		41	44.57
(4) 4 ครั้ง		7	7.61
(5) 5 ครั้ง		2	2.17
	รวม	92	100
ทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือวัด			
(1) CTT		78	84.78
(2) IRT		1	1.09
(3) CCT & G Theory		4	4.35
(4) CCT & IRT		8	8.70
(5) G Theory & IRT		1	1.09
	รวม	92	100
การตรวจให้คะแนน			
(1) ให้คะแนนแบบสองค่า		17	18.48
(2) ให้คะแนนมากกว่าสองค่า		52	56.52
(3) ให้คะแนนแบบสองค่า และมากกว่าสองค่า		23	25.00
	รวม	92	100

ตัวแปรคุณลักษณะของงานวิจัย	ความถี่	ร้อยละ	
ประเภทของเกณฑ์ปกติ			
(1) เกณฑ์ปกติภายในชั้นเรียน	1	1.09	
(2) เกณฑ์ปกติระดับสถาบัน	1	1.09	
(3) เกณฑ์ปกติระดับท้องถิ่น	33	35.87	
(4) เกณฑ์ปกติระดับชาติ	1	1.09	
(5) เกณฑ์ปกติระดับท้องถิ่น และเกณฑ์ปกติระดับชาติ	1	1.09	
(6) ไม่ระบุ	25	27.17	
(7) ไม่มีการสร้างเกณฑ์ปกติ	30	32.61	
รวม	92	100	
วิธีที่ใช้ในการสร้างเกณฑ์ปกติ			
(1) วิธีเปอร์เซ็นต์ไทล์	3	3.26	
(2) วิธีคะแนนมาตรฐานที่	1	1.09	
(3) วิธีคะแนนมาตรฐานที่ปกติ	54	58.70	
(4) วิธีเปอร์เซ็นต์ไทล์และคะแนนมาตรฐานที่รวมกัน	1	1.09	
(5) วิธีเปอร์เซ็นต์ไทล์และสเตรโนนรวมกัน	3	3.26	
(6) ไม่มีการสร้างเกณฑ์ปกติ	30	32.61	
รวม	92	100	
คู่มือการใช้เครื่องมือวัด	มี	70	76.09
	ไม่มี	22	23.91
	รวม	92	100
ด้านวิธีวิทยาการวิจัย			
การกำหนดขนาดของตัวอย่าง			
(1) สูตรของ Krejcie & Morgan (1970)	8	8.70	
(2) สูตรของ Yamane (1973)	20	21.74	
(3) สูตรของ Darwin	3	3.26	
(4) สูตรคำนวณของบุญชม ศรีสะอาด (2553)	2	2.17	
(5) เปิดตารางของ ศิริชัย กาญจนวาสี และคณะ (2535)	2	2.17	
(6) กำหนดขนาดตัวอย่างตามสัดส่วน	6	6.52	
(7) เกณฑ์ร้อยละของประชากร	4	4.35	
(8) สูตรกำหนดขนาดด้วยการสุ่มแบบแบ่งชั้น	2	2.17	
(9) เกณฑ์กำหนดขนาดตัวอย่างสำหรับพัฒนาเครื่องมือ ของ บุญชม ศรีสะอาด (2534)	3	3.26	
(10) โปรแกรม G* Power	3	3.26	
(11) 10 เท่าของตัวแปร/พารามิเตอร์	5	5.43	
(12) 15 เท่าของตัวแปร/พารามิเตอร์	1	1.09	
(13) 20 เท่าของตัวแปร/พารามิเตอร์	14	15.22	
(14) หลาก ๆ เกณฑ์	9	9.78	
(15) อื่น ๆ	3	3.26	
(16) ไม่ระบุ	7	7.61	
รวม	92	100	

ตัวแปรคุณลักษณะของงานวิจัย		ความถี่	ร้อยละ
วิธีการเลือกตัวอย่าง			
(1) สุ่มอย่างง่าย		2	2.17
(2) สุ่มแบบกลุ่ม		1	1.09
(3) สุ่มแบบแบ่งชั้น		12	13.04
(4) สุ่มแบบสองขั้นตอน		20	21.74
(5) สุ่มแบบหลายขั้นตอน		46	50.00
(6) เลือกอย่างเจาะจง		5	5.43
(7) ไม่ระบุ		6	6.52
	รวม	92	100
การวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย			
(1) content analysis	ใช่	7	7.61
	ไม่ใช่	85	92.39
	รวม	92	100
(2) descriptive stat	ใช่	67	72.83
	ไม่ใช่	25	27.17
	รวม	92	100
(3) independent t-test	ใช่	7	7.61
	ไม่ใช่	85	92.39
	รวม	92	100
(4) correlation	ใช่	6	6.52
	ไม่ใช่	86	93.48
	รวม	92	100
(5) one-way ANOVA	ใช่	5	5.43
	ไม่ใช่	87	94.57
	รวม	92	100
(6) two-way ANOVA	ใช่	4	4.35
	ไม่ใช่	88	95.65
	รวม	92	100
(7) three-way ANOVA	ใช่	2	2.17
	ไม่ใช่	90	97.83
	รวม	92	100
(8) MANOVA	ใช่	1	1.09
	ไม่ใช่	91	98.91
	รวม	92	100
(9) exploratory factor analysis	ใช่	4	4.35
	ไม่ใช่	88	95.65
	รวม	92	100
(10) confirmatory factor analysis	ใช่	44	47.83
	ไม่ใช่	48	52.17
	รวม	92	100

ตัวแปรคุณลักษณะของงานวิจัย		ความถี่	ร้อยละ
(11) ทดสอบความไม่แปรเปลี่ยน	ใช่	4	4.35
	ไม่ใช่	88	95.65
	รวม	92	100
(12) ค่าความเที่ยง	ใช่	92	100
	ไม่ใช่	0	0
	รวม	92	100
(13) ค่าความตรง	ใช่	92	100
	ไม่ใช่	0	0
	รวม	92	100
(14) ค่าความยาก	ใช่	51	55.43
	ไม่ใช่	41	44.57
	รวม	92	100
(15) ค่าอำนาจจำแนก	ใช่	89.0	96.7
	ไม่ใช่	3.0	3.3
	รวม	92	100
ด้านคุณภาพงานวิจัย			
(1) ดีมาก		16	17.39
(2) ดี		76	82.61
	รวม	92	100

* งานวิจัย 1 เรื่องมีการศึกษามากกว่า 1 รายการ

2.2 ผลการวิเคราะห์ข้อมูลเบื้องต้นจำแนกตามตัวแปรคุณลักษณะงานวิจัยที่เป็นตัวแปรต่อเนื่อง

ตัวแปรคุณลักษณะของงานวิจัยที่เป็นตัวแปรต่อเนื่อง ประกอบด้วย ตัวแปรคุณลักษณะของงานวิจัยด้านเนื้อหาสาระงานวิจัย และด้านวิธีวิทยาการวิจัย จำนวน 11 ตัวแปร จากงานวิจัยที่นำมาสังเคราะห์จำนวน 92 เล่ม มีผลการวิเคราะห์ดังนี้

เมื่อพิจารณาตัวแปรคุณลักษณะของงานวิจัยด้านเนื้อหาสาระงานวิจัย ประกอบด้วย ตัวแปรต่อเนื่องจำนวน 5 ตัวแปร คือ จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา, จำนวนผู้เชี่ยวชาญด้านการวัดผล, จำนวนผู้เชี่ยวชาญด้านการเนื้อหา, จำนวนข้อที่สร้างก่อนการทดลองใช้ และจำนวนข้อคำถามใช้จริง ผลการวิเคราะห์พบว่า ตัวแปรจำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา มีค่าเฉลี่ย (M) เท่ากับ 6.217 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 3.454 และมีค่าสัมประสิทธิ์การกระจาย ($C.V.$) เท่ากับ 55.557 เมื่อพิจารณาการแจกแจงของข้อมูล พบว่า การแจกแจงของข้อมูลมีลักษณะเบ้ขวา ($Sk = 2.920$) แสดงว่า จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหามีค่าต่ำกว่าค่าเฉลี่ย และมีการแจกแจงของข้อมูลสูงกว่าโค้งปกติ ($Ku = 8.375$) แสดงว่าข้อมูลมีการกระจายน้อย ตัวแปรจำนวนผู้เชี่ยวชาญด้านการวัดผลมี ค่าเฉลี่ย (M) เท่ากับ 2.360 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 1.579 และมีค่าสัมประสิทธิ์การกระจาย ($C.V.$) เท่ากับ 66.907 เมื่อพิจารณาการแจกแจงของข้อมูล พบว่า การแจกแจงของข้อมูลมีลักษณะเบ้ขวา ($Sk = 3.276$) แสดงว่า จำนวนผู้เชี่ยวชาญด้านการวัดผลมีค่าต่ำกว่าค่าเฉลี่ย และมีการแจกแจงของข้อมูลสูงกว่าโค้งปกติ ($Ku = 17.045$) แสดงว่าข้อมูลมีการกระจายน้อย ตัวแปรจำนวนผู้เชี่ยวชาญด้านการเนื้อหา มีค่าเฉลี่ย (M) เท่ากับ 3.663 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 2.379 และมีค่าสัมประสิทธิ์การกระจาย ($C.V.$) เท่ากับ 64.947 เมื่อพิจารณาการแจกแจงของข้อมูล พบว่า

ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ 258.094 และมีค่าสัมประสิทธิ์การกระจาย ($C.V.$) เท่ากับ 51.878 รายละเอียดดังตารางที่ 4.4

ตารางที่ 4.4 ข้อมูลพื้นฐานของตัวแปรคุณลักษณะงานวิจัยที่เป็นตัวแปรต่อเนื่อง

ตัวแปรคุณลักษณะของงานวิจัย	n	M	SD	C.V.	min	max	Sk	Ku
1. จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา	92	6.217	3.454	55.557	3	20	2.920	8.375
2. จำนวนผู้เชี่ยวชาญด้านการวัดผล	86	2.360	1.579	66.907	0	12	3.276	17.045
3. จำนวนผู้เชี่ยวชาญด้านการเนื้อหา	86	3.663	2.379	64.947	0	15	2.896	11.547
4. จำนวนข้อที่สร้างก่อนการทดลองใช้	91	54.780	104.941	191.568	3	1000	8.312	75.108
5. จำนวนข้อคำถามใช้จริง	92	37.500	66.479	177.277	4	622	7.739	67.286
6. ขนาดของตัวอย่างทดลองใช้เครื่องมือ (ภาพรวม)	92	866.446	792.812	91.502	40	6300	3.868	23.717
7. ขนาดของตัวอย่างทดลองใช้เครื่องมือ ครั้งที่ 1	92	130.326	156.619	120.175	3	972	3.273	12.788
8. ขนาดของตัวอย่างทดลองใช้เครื่องมือ ครั้งที่ 2	90	413.033	735.573	178.091	15	6000	5.431	37.857
9. ขนาดของตัวอย่างทดลองใช้เครื่องมือ ครั้งที่ 3	51	496.627	420.969	84.766	20	2532	2.550	10.010
10. ขนาดของตัวอย่างทดลองใช้เครื่องมือ ครั้งที่ 4	9	294.556	233.233	79.181	36	584	0.105	-2.021
11. ขนาดของตัวอย่างทดลองใช้เครื่องมือ ครั้งที่ 5	2	497.500	258.094	51.878	315	680	-	-

2.3 ผลการวิเคราะห์ข้อมูลเบื้องต้นของการทดลองใช้เครื่องมือ

ผลการวิเคราะห์ข้อมูลเบื้องต้นเกี่ยวกับการทดลองใช้เครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ซึ่งเป็นการนำเสนอผลการนำเครื่องมือไปใช้เพื่อการตรวจสอบคุณภาพและการตอบวัตถุประสงค์ในการวิจัย ผลการวิเคราะห์มีดังนี้

การทดลองใช้เครื่องมือครั้งที่ 1 งานวิจัยทุกเรื่องมีการทดลองใช้ (ร้อยละ 100) โดยมีวัตถุประสงค์ของการทดลองใช้ในครั้งนี้ 1 เพื่อตรวจสอบค่าชี้แจง/วิธีดำเนินการสอบมากที่สุด (ร้อยละ 86.96) รองลงมาคือ ตรวจสอบอำนาจจำแนก (ร้อยละ 80.43), ตรวจสอบความยาก (ร้อยละ 43.48), ตรวจสอบเวลาที่ใช้ (ร้อยละ 38.04), ตรวจสอบภาษา (ร้อยละ 36.96), ตรวจสอบค่าความเที่ยง (ร้อยละ 28.26), ตรวจสอบตัวเลือก (ร้อยละ 6.52) และตรวจสอบค่าความตรง (ร้อยละ 3.26) ตามลำดับ โดยงานวิจัยส่วนใหญ่ไม่ได้ระบุวิธีการเลือกตัวอย่างในการทดลองใช้ (ร้อยละ 47.83) รองลงมา 3 อันดับแรก ใช้การสุ่มแบบหลายขั้นตอน (ร้อยละ 17.39) เลือกอย่างเจาะจง (ร้อยละ 11.96), และสุ่มแบบสองขั้นตอน (ร้อยละ 10.87) ตามลำดับ งานวิจัยส่วนใหญ่แสดงดัชนีคุณภาพ (ร้อยละ 70.65) สำหรับการประมาณค่าความเที่ยงงานวิจัยมีการใช้ Cronbach's alpha method มากที่สุด (ร้อยละ 20.65) รองลงมาคือ ใช้มากกว่า 1 วิธี (ร้อยละ 5.43) KR20/KR21 และ Lovett (ร้อยละ 1.09) ตามลำดับ ในขณะที่การตรวจสอบความตรงมีเพียงการตรวจสอบความตรงตามสภาพ (ร้อยละ 2.17) และความตรงตามโครงสร้างด้วยการวิเคราะห์องค์ประกอบ (ร้อยละ 1.09) สำหรับการตรวจสอบความยาก งานวิจัยใช้สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำมากที่สุด (ร้อยละ 16.30) รองลงมา 3 อันดับแรก คือ สูตร Whitney and Sabers และใช้มากกว่า 1 วิธี มีจำนวนเท่ากัน (ร้อยละ 6.52) และวิธีอื่น ๆ (ร้อยละ 5.43) นอกจากนี้ งานวิจัยมีการหาค่าอำนาจจำแนกด้วยสถิติทดสอบที่

(t-test) มากที่สุด (ร้อยละ 18.48) รองลงมา 3 อันดับแรก คือ item-total correlation (ร้อยละ 17.39), มากกว่า 1 วิธี (ร้อยละ 11.96) และสูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ (ร้อยละ 9.78) ตามลำดับ

การทดลองใช้เครื่องมือครั้งที่ 2 งานวิจัยส่วนใหญ่มีการทดลองใช้ (ร้อยละ 97.83) โดยมีวัตถุประสงค์ของการทดลองใช้ในครั้งที่ 2 เพื่อตรวจสอบอำนาจจำแนกมากที่สุด (ร้อยละ 71.74) รองลงมา 3 อันดับแรก คือ ตรวจสอบค่าความเที่ยง (ร้อยละ 60.87) ตรวจสอบความยาก (ร้อยละ 43.48) และตรวจสอบค่าความตรง (ร้อยละ 35.87) ตามลำดับ โดยมีงานวิจัยใช้วิธีสุ่มแบบหลายขั้นตอน และไม่มีระบุวิธีสุ่มมีจำนวนเท่ากัน (ร้อยละ 27.17) รองลงมา 3 อันดับแรก ใช้การสุ่มแบบสองขั้นตอน (ร้อยละ 17.39), เลือกอย่างเจาะจง (ร้อยละ 9.78) และสุ่มแบบแบ่งชั้น (ร้อยละ 7.61) ตามลำดับ งานวิจัยส่วนใหญ่แสดงดัชนีคุณภาพ (ร้อยละ 76.09) สำหรับการประมาณค่าความเที่ยงงานวิจัยมีการใช้ Cronbach's alpha method มากที่สุด (ร้อยละ 31.52) รองลงมาคือ ใช้มากกว่า 1 วิธี (ร้อยละ 17.39), KR20/KR21 (ร้อยละ 8.70), Hoyt's analysis (ร้อยละ 2.17) และ G-theory (ร้อยละ 1.09) ตามลำดับ ในขณะที่การตรวจสอบความตรงพบว่า มีการตรวจสอบความตรงตามโครงสร้างด้วยการวิเคราะห์องค์ประกอบ มากที่สุด (ร้อยละ 23.91) รองลงมา คือ ความตรงตามโครงสร้างด้วยวิธี known group (ร้อยละ 4.35), ความตรงตามสภาพ (ร้อยละ 3.26) และความตรงตามโครงสร้างด้วยการวิเคราะห์เมทริกซ์พหุลักษณะวิธี (ร้อยละ 1.09) ตามลำดับ สำหรับการตรวจสอบความยาก งานวิจัยใช้สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำมากที่สุด (ร้อยละ 14.13) รองลงมา 3 อันดับแรก คือ สูตร Whitney and Sabers (ร้อยละ 7.61), ใช้มากกว่า 1 วิธี (ร้อยละ 6.52) และ IRT (ร้อยละ 4.35) ตามลำดับ นอกจากนี้ งานวิจัยมีการหาค่าอำนาจจำแนกด้วย item-total correlation และมากกว่า 1 วิธี มากที่สุด (ร้อยละ 15.22) รองลงมา 3 อันดับแรก คือ สถิติทดสอบที (t-test) และสูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ มีจำนวนเท่ากัน (ร้อยละ 8.70) และ Point Biserial correlation (ร้อยละ 5.43) ตามลำดับ

การทดลองใช้เครื่องมือครั้งที่ 3 งานวิจัยหนึ่งมีการทดลองใช้ (ร้อยละ 55.43) โดยมีวัตถุประสงค์ของการทดลองใช้ในครั้งที่ 3 เพื่อตรวจสอบค่าความเที่ยง (ร้อยละ 45.65) รองลงมา 3 อันดับแรก คือ ตรวจสอบค่าความตรง (ร้อยละ 36.96), ตรวจสอบอำนาจจำแนก (ร้อยละ 25.00) และตรวจสอบความยาก (ร้อยละ 17.39) ตามลำดับ โดยมีงานวิจัยใช้วิธีสุ่มแบบหลายขั้นตอน มากที่สุด (ร้อยละ 20.65) รองลงมา 3 อันดับแรก ใช้การสุ่มแบบสองขั้นตอน (ร้อยละ 10.87), สุ่มแบบแบ่งชั้น (ร้อยละ 9.78) และใช้ประชากร/กลุ่มเป้าหมาย (ร้อยละ 4.35) ตามลำดับ งานวิจัยแสดงดัชนีคุณภาพ (ร้อยละ 38.04) สำหรับการประมาณค่าความเที่ยงงานวิจัยมีการใช้ Cronbach's alpha method มากที่สุด (ร้อยละ 29.35) รองลงมาคือ ใช้มากกว่า 1 วิธี (ร้อยละ 8.70), KR20/KR21 (ร้อยละ 5.43) Lovett และ G-theory จำนวนเท่ากัน (ร้อยละ 1.09) ตามลำดับ ในขณะที่การตรวจสอบความตรงพบว่า มีการตรวจสอบความตรงตามโครงสร้างด้วยการวิเคราะห์องค์ประกอบ มากที่สุด (ร้อยละ 18.48) รองลงมา 3 อันดับแรก คือ ใช้มากกว่า 1 วิธี (ร้อยละ 5.43), ความตรงตามโครงสร้างด้วยวิธี known group (ร้อยละ 4.35) และความตรงตามสภาพ (ร้อยละ 3.26) ตามลำดับ สำหรับการตรวจสอบความยาก งานวิจัยใช้สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำมากที่สุด (ร้อยละ 7.61) รองลงมา 3 อันดับแรก คือ อื่น ๆ (ร้อยละ 3.26), สูตร Whitney and Sabers (ร้อยละ 2.71), และใช้โปรแกรม TAP, IRT, ใช้มากกว่า 1 วิธี และไม่มีระบุ มีจำนวนเท่ากัน (ร้อยละ 1.09) ตามลำดับ นอกจากนี้ งานวิจัยมีการหาค่าอำนาจจำแนกด้วย item-total correlation มากที่สุด (ร้อยละ 5.43) รองลงมา 3 อันดับแรก คือ สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ และมากกว่า 1 วิธี มีจำนวนเท่ากัน (ร้อยละ 4.35) และ สถิติทดสอบที (t-test), Point Biserial correlation, โปรแกรม TAP และอื่น ๆ มีจำนวนเท่ากัน (ร้อยละ 2.17) ตามลำดับ

การทดลองใช้เครื่องมือครั้งที่ 4 งานวิจัยหนึ่งมีการทดลองใช้ (ร้อยละ 9.78) โดยมีวัตถุประสงค์ของการทดลองใช้ในครั้งที่ 4 เพื่อตรวจสอบค่าความตรง (ร้อยละ 7.61) รองลงมาคือ ตรวจสอบค่าความเที่ยง (ร้อยละ 5.43), ตรวจสอบความยาก และตรวจสอบอำนาจจำแนก มีจำนวนเท่ากัน (ร้อยละ 3.26) โดยมีงานวิจัยไม่ระบุวิธีสุ่ม มากที่สุด (ร้อยละ 3.26) รองลงมา คือ สุ่มแบบสองขั้นตอน และสุ่มแบบหลายขั้นตอน มีจำนวนเท่ากัน (ร้อยละ 2.17), สุ่มแบบแบ่งชั้น และใช้ประชากร/กลุ่มเป้าหมาย มีจำนวนเท่ากัน (ร้อยละ 1.09) ตามลำดับ งานวิจัยแสดงดัชนีคุณภาพ (ร้อยละ 5.43) สำหรับการประมาณค่าความเที่ยงงานวิจัยมีการใช้ KR20/KR21 และ Cronbach's alpha method มากที่สุด (ร้อยละ 2.17) รองลงมาคือ test-retest method (ร้อยละ 1.09) ในขณะที่การตรวจสอบความตรง พบว่า มีการตรวจสอบความตรงตามสภาพ มากที่สุด (ร้อยละ 3.26) รองลงมาคือ ความตรงตามโครงสร้างด้วยการวิเคราะห์องค์ประกอบ (ร้อยละ 2.17), ความตรงตามโครงสร้างด้วยวิธี known group และใช้มากกว่า 1 วิธี (ร้อยละ 1.09) สำหรับการตรวจสอบความยาก งานวิจัยใช้สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำมากที่สุด (ร้อยละ 2.17) รองลงมาคือ ไม่ระบุ (ร้อยละ 1.09) ตามลำดับ สำหรับการตรวจสอบอำนาจจำแนก งานวิจัยใช้ สถิติทดสอบที (t-test), Point Biserial correlation และโปรแกรม TAP มีจำนวนเท่ากัน (ร้อยละ 1.09)

การทดลองใช้เครื่องมือครั้งที่ 5 งานวิจัยหนึ่งมีการทดลองใช้ (ร้อยละ 2.17) โดยมีวัตถุประสงค์ของการทดลองใช้ในครั้งที่ 5 คือ ตรวจสอบค่าความเที่ยง (ร้อยละ 2.17) โดยมีงานวิจัยไม่ระบุวิธีสุ่ม (ร้อยละ 2.17) สำหรับการประมาณค่าความเที่ยงงานวิจัยมีการใช้ KR20/KR21 และ Cronbach's alpha method จำนวนเท่ากัน (ร้อยละ 1.09) ทั้งนี้ไม่มีการตรวจสอบความตรง ความยาก และอำนาจจำแนก รายละเอียดดังตารางที่ 4.5

ตารางที่ 4.5 ข้อมูลเบื้องต้นของการทดลองใช้เครื่องมือ

ตัวแปรคุณลักษณะ ของงานวิจัย	ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ครั้งที่ 4		ครั้งที่ 5		
	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ	
การทดลองใช้เครื่องมือ	92	100	90	97.83	51	55.43	9	9.78	2	2.17	
วัตถุประสงค์ของการทดลองใช้											
(1) ตรวจสอบค่าความ เที่ยง	ใช่	26	28.26	56	60.87	42	45.65	5	5.43	2	2.17
	ไม่ใช่	66	71.74	36	39.13	50	54.35	87	94.57	90	97.83
	รวม	92	100	92	100	92	100	92	100	92	100
(2) ตรวจสอบค่าความ ตรง	ใช่	3	3.26	33	35.87	34	36.96	7	7.61	0	0
	ไม่ใช่	89	96.74	59	64.13	58	63.04	85	92.39	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
(3) ตรวจสอบความยาก	ใช่	40	43.48	37	40.22	16	17.39	3	3.26	0	0
	ไม่ใช่	52	56.52	55	59.78	76	82.61	89	96.74	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
(4) ตรวจสอบอำนาจ จำแนก	ใช่	74	80.43	66	71.74	23	25.00	3	3.26	0	0
	ไม่ใช่	18	19.57	26	28.26	69	75.00	89	96.74	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
(5) ตรวจสอบภาษา	ใช่	34	36.96	3	3.26	0	0	0	0	0	0
	ไม่ใช่	58	63.04	89	96.74	92	100	92	100	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
(6) ตรวจสอบเวลาที่ใช้	ใช่	35	38.04	2	2.17	1	1.09	0	0	0	0
	ไม่ใช่	57	61.96	90	97.83	91	98.91	92	100	92	100
	รวม	92	100	92	100	92	100	92	100	92	100

ตัวแปรคุณลักษณะ ของงานวิจัย		ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ครั้งที่ 4		ครั้งที่ 5	
		ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ
(7) ตรวจสอบตัวเลือก	ใช่	6	6.52	1	1.09	0	0	0	0	0	0
	ไม่ใช่	86	93.48	91	98.91	92	100	92	100	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
(8) ตรวจสอบความ ครอบคลุมของเนื้อหา	ใช่	0	0	1	1.09	0	0	0	0	0	0
	ไม่ใช่	92	100	91	98.91	92	100	92	100	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
(9) ตรวจสอบคำชี้แจง/ วิธีดำเนินการสอบ	ใช่	80	86.96	1	1.09	0	0	0	0	0	0
	ไม่ใช่	80	86.96	91	98.91	92	100	92	100	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
(10) ตรวจสอบความไม่ แปรเปลี่ยน	ใช่	0	0	3	3.26	1	1.09	0	0	0	0
	ไม่ใช่	92	100	89	96.74	91	98.91	92	100	92	100
	รวม	92	100	92	100	92	100	92	100	92	100
วิธีการเลือกตัวอย่างในการทดลองใช้											
(1) ใช้ประชากร/กลุ่มเป้าหมาย		1	1.09	2	2.17	4	4.35	1	1.09	0	0
(2) สุ่มอย่างง่าย		5	5.43	5	5.43	3	3.26	0	0	0	0
(3) สุ่มแบบกลุ่ม		1	1.09	1	1.09	1	1.09	0	0	0	0
(4) สุ่มแบบแบ่งชั้น		4	4.35	7	7.61	9	9.78	1	1.09	0	0
(5) สุ่มแบบสองขั้นตอน		10	10.87	16	17.39	10	10.87	2	2.17	0	0
(6) สุ่มแบบหลายขั้นตอน		16	17.39	25	27.17	19	20.65	2	2.17	0	0
(7) เลือกอย่างเจาะจง		11	11.96	9	9.78	2	2.17	0	0	0	0
(8) ไม่ระบุ		44	47.83	25	27.17	3	3.26	3	3.26	2	2.17
(9) ไม่มีการทดลองใช้		0	0	2	2.17	41	44.57	83	90.2	90	97.83
รวม		92	100	92	100	92	100	92	100	92	100
ดัชนีแสดงคุณภาพ	มี	65	70.65	70	76.09	35	38.04	5	5.43	0	0
	ไม่มี	27	29.35	20	21.74	16	17.39	4	4.35	2	2.17
	ไม่ทดลอง	0	0	2	2.17	41	44.57	83	90.22	90	97.83
	รวม	92	100	92	100	92	100	92	100	92	100
ประเภทความเที่ยง											
(1) test-retest method		0	0	0	0	0	0	1	1.09	0	0
(2) KR20/KR21		1	1.09	8	8.70	5	5.43	2	2.17	1	1.09
(3) Cronbach's alpha method		19	20.65	29	31.52	27	29.35	2	2.17	1	1.09
(4) Lovett		1	1.09	0	0	1	1.09	0	0	0	0
(5) Hoyt's analysis		0	0	2	2.17	0	0	0	0	0	0
(6) G-theory		0	0	1	1.09	1	1.09	0	0	0	0
(7) ใช้มากกว่า 1 วิธี		5	5.43	16	17.39	8	8.70	0	0	0	0
(8) ไม่มีการตรวจสอบ		66	71.74	36	39.13	50	54.35	87	94.57	90	97.83
รวม		92	100	92	100	92	100	92	100	92	100
ประเภทความตรง											
(1) ความตรงตามสภาพ		2	2.17	3	3.26	3	3.26	3	3.26	0	0
(2) ความตรงตามโครงสร้างด้วย การตัดสินโดยผู้เชี่ยวชาญ		0	0	0	0	1	1.09	0	0	0	0
(3) ความตรงตามโครงสร้างด้วยวิธี known group		0	0	4	4.35	4	4.35	1	1.09	0	0
(4) ความตรงตามโครงสร้างด้วยการ วิเคราะห์เมทริกซ์พหุลักษณะวิธี		0	0	1	1.09	1	1.09	0	0	0	0
(5) ความตรงตามโครงสร้างด้วย การวิเคราะห์องค์ประกอบ		1	1.09	22	23.91	17	18.48	2	2.17	0	0

ตัวแปรคุณลักษณะ ของงานวิจัย	ครั้งที่ 1		ครั้งที่ 2		ครั้งที่ 3		ครั้งที่ 4		ครั้งที่ 5	
	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ	ความถี่	ร้อยละ
(6) ใช้มากกว่า 1 วิธี	0	0	0	0	5	5.43	1	1.09	0	0
(7) ไม่มีการตรวจสอบ	89	96.74	3	3.26	58	63.04	85	92.39	92	100
รวม	92	100	92	100	92	100	92	100	92	100
วิธีการหาค่าความยาก										
สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ	15	16.30	13	14.13	7	7.61	2	2.17	0	0
สูตร Whitney and Sabers	6	6.52	7	7.61	2	2.17	0	0	0	0
โปรแกรม Evana	2	2.17	0	0	0	0	0	0	0	0
โปรแกรม B-index	2	2.17	1	1.09	0	0	0	0	0	0
โปรแกรม TAP	2	2.17	1	1.09	1	1.09	0	0	0	0
IRT	1	1.09	4	4.35	1	1.09	0	0	0	0
อื่น ๆ	5	5.43	3	3.26	3	3.26	0	0	0	0
ใช้มากกว่า 1 วิธี	6	6.52	6	6.52	1	1.09	0	0	0	0
ไม่ระบุ	1	1.09	2	2.17	1	1.09	1	1.09	0	0
ไม่มีการตรวจสอบ	85	92.39	55	59.78	76	82.61	89	96.7	92	100
รวม	92	100	92	100	92	100	92	100	92	100
วิธีการหาค่าอำนาจจำแนก										
สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ	9	9.78	8	8.70	4	4.35	1	1.09	0	0
item-total correlation	16	17.39	14	15.22	5	5.43	0	0	0	0
สถิติทดสอบที (t-test)	17	18.48	8	8.70	2	2.17	0	0	0	0
Biserial correlation	1	1.09	1	1.09	0	0	0	0	0	0
Point Biserial correlation	7	7.61	5	5.43	2	2.17	1	1.09	0	0
Whitney and Sabers	4	4.35	4	4.35	1	1.09	0	0	0	0
โปรแกรม B-index	3	3.26	2	2.17	0	0	0	0	0	0
โปรแกรม TAP	1	1.09	2	2.17	2	2.17	1	1.09	0	0
IRT	0	0	4	4.35	1	1.09	0	0	0	0
อื่น ๆ	5	5.43	4	4.35	2	2.17	0	0	0	0
มากกว่า 1 วิธี	11	11.96	14	15.22	4	4.35	0	0	0	0
ไม่มีการตรวจสอบ	18	19.57	26	28.26	69	75.00	89	96.74	0	0
รวม	92	100	92	100	92	100	92	100	92	100

ตอนที่ 3 ผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

การวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ผู้วิจัยนำงานวิจัยที่มีการตรวจสอบคุณภาพเครื่องมือตามทฤษฎีการทดสอบแบบดั้งเดิมมาใช้ในการวิเคราะห์หोगิมาณ ผลจากการสำรวจพบว่าม้งานวิจัยจำนวน 90 เล่ม จากจำนวนทั้งหมด 92 เล่ม ที่เป็นไปตามเกณฑ์ดังกล่าว สำหรับงานวิจัยอีก 2 เล่ม ผู้วิจัยได้นำไปใช้ในการสังเคราะห์ข้อมูลเชิงคุณภาพต่อไป สำหรับการนำเสนอในตอนที 3 นี้ ผู้วิจัยแบ่งการนำเสนอออกเป็น 2 ส่วน คือ 3.1) การวิเคราะห์ข้อมูลเบื้องต้นของดัชนีมาตรฐาน และ 3.2) การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานจำแนกตามตัวแปรคุณลักษณะงานวิจัย มีรายละเอียดดังนี้

3.1 การวิเคราะห์ข้อมูลเบื้องต้นของค่าดัชนีมาตรฐาน

ค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษาที่ใช้ในการวิจัยครั้งนี้มีทั้งหมด 4 ค่า ได้แก่ ค่าดัชนีมาตรฐานความเที่ยง ค่าดัชนีมาตรฐานความตรง

ค่าดัชนีมาตรฐานความยาก และค่าดัชนีมาตรฐานอำนาจจำแนก มีรายละเอียดของผลการวิเคราะห์ข้อมูลเบื้องต้นของค่าดัชนีมาตรฐานแต่ละค่า ดังนี้

ค่าดัชนีมาตรฐานความเที่ยงที่ใช้ในการวิจัยครั้งนี้มีจำนวน 216 ค่า ผลการวิเคราะห์ข้อมูลเบื้องต้น พบว่า ค่าเฉลี่ยดัชนีมาตรฐานความเที่ยง (M) เท่ากับ .778 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .163 และมีค่าสัมประสิทธิ์การกระจาย (C.V.) เท่ากับ 20.951 ค่าดัชนีมาตรฐานความเที่ยงที่มีค่าต่ำสุดมีค่าเท่ากับ .130 ค่าดัชนีมาตรฐานความเที่ยงมีค่าสูงสุดมีค่าเท่ากับ .999 เมื่อพิจารณาการแจกแจงของข้อมูลพบว่า การแจกแจงของข้อมูลมีลักษณะเบ้ซ้าย ($Sk = -1.346$) แสดงว่า ค่าดัชนีมาตรฐานความเที่ยงมีค่าสูงกว่าค่าเฉลี่ย และมีการแจกแจงของข้อมูลสูงกว่าโค้งปกติ ($Ku = 1.968$) แสดงว่าข้อมูลมีการกระจายน้อย

ค่าดัชนีมาตรฐานความตรงที่ใช้ในการวิจัยครั้งนี้มีจำนวน 39 ค่า ผลการวิเคราะห์ข้อมูลเบื้องต้น พบว่า ค่าเฉลี่ยดัชนีมาตรฐานความตรง (M) เท่ากับ .596 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .199 และมีค่าสัมประสิทธิ์การกระจาย (C.V.) เท่ากับ 33.389 ค่าดัชนีมาตรฐานความตรงที่มีค่าต่ำสุดมีค่าเท่ากับ .240 ค่าดัชนีมาตรฐานความตรงมีค่าสูงสุดมีค่าเท่ากับ .995 เมื่อพิจารณาการแจกแจงของข้อมูลพบว่า การแจกแจงของข้อมูลมีลักษณะเบ้ขวา ($Sk = .538$) แสดงว่า ค่าดัชนีมาตรฐานความตรงมีค่าต่ำกว่าค่าเฉลี่ย และมีการแจกแจงของข้อมูลต่ำกว่าโค้งปกติ ($Ku = -.502$) แสดงว่าข้อมูลมีการกระจายมาก

ค่าดัชนีมาตรฐานความยากที่ใช้ในการวิจัยครั้งนี้มีจำนวน 94 ค่า ผลการวิเคราะห์ข้อมูลเบื้องต้น พบว่า ค่าเฉลี่ยดัชนีมาตรฐานความยาก (M) เท่ากับ .537 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .113 และมีค่าสัมประสิทธิ์การกระจาย (C.V.) เท่ากับ 21.043 ค่าดัชนีมาตรฐานความยากที่มีค่าต่ำสุดมีค่าเท่ากับ .230 ค่าดัชนีมาตรฐานความยากมีค่าสูงสุดมีค่าเท่ากับ .815 เมื่อพิจารณาการแจกแจงของข้อมูลพบว่า การแจกแจงของข้อมูลมีลักษณะเบ้ซ้ายเล็กน้อย ($Sk = -.051$) แสดงว่า ค่าดัชนีมาตรฐานความยากมีค่าสูงกว่าค่าเฉลี่ยเล็กน้อย และมีการแจกแจงของข้อมูลสูงกว่าโค้งปกติเล็กน้อย ($Ku = .184$) แสดงว่าข้อมูลมีการกระจายค่อนข้างน้อย

ค่าดัชนีมาตรฐานอำนาจจำแนกที่ใช้ในการวิจัยครั้งนี้มีจำนวน 161 ค่า ผลการวิเคราะห์ข้อมูลเบื้องต้น พบว่า ค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนก (M) เท่ากับ .460 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .137 และมีค่าสัมประสิทธิ์การกระจาย (C.V.) เท่ากับ 29.783 ค่าดัชนีมาตรฐานอำนาจจำแนกที่มีค่าต่ำสุดมีค่าเท่ากับ .194 ค่าดัชนีมาตรฐานอำนาจจำแนกมีค่าสูงสุดมีค่าเท่ากับ .823 เมื่อพิจารณาการแจกแจงของข้อมูลพบว่า การแจกแจงของข้อมูลมีลักษณะเบ้ขวา ($Sk = .254$) แสดงว่า ค่าดัชนีมาตรฐานอำนาจจำแนกมีค่าต่ำกว่าค่าเฉลี่ย และมีการแจกแจงของข้อมูลต่ำกว่าโค้งปกติ ($Ku = -.507$) แสดงว่าข้อมูลมีการกระจายมาก รายละเอียดดังตารางที่ 4.6

ตารางที่ 4.6 ผลการวิเคราะห์ข้อมูลเบื้องต้นของค่าดัชนีมาตรฐาน

ค่าดัชนีมาตรฐาน	n	M	SD	C.V.	min	max	Sk	Ku
ค่าความเที่ยง	216	.778	.163	20.951	.130	.999	-1.346	1.968
ค่าความตรง	39	.596	.199	33.389	.240	.995	.538	-.502
ค่าความยาก	94	.537	.113	21.043	.230	.815	-.051	.184
ค่าอำนาจจำแนก	161	.460	.137	29.783	.194	.823	.254	-.507

3.2 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานจำแนกตามตัวแปรคุณลักษณะงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานจำแนกตามตัวแปรคุณลักษณะงานวิจัย ด้วยการวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA) และทดสอบเอกพันธ์ของความแปรปรวน (test of homogeneity of variance) ตรวจสอบข้อตกลงเบื้องต้นโดยการใช้การทดสอบ Levene (Levene's test) เพื่อทำการทดสอบความแตกต่างของค่าเฉลี่ยรายคู่ ในกรณีที่ความแปรปรวนมีความเป็นเอกพันธ์ผู้วิจัยใช้การทดสอบรายคู่ภายหลังด้วย Bonferroni และใช้ Dunnett's T3 ในกรณีที่ความแปรปรวนไม่เป็นเอกพันธ์ ทั้งนี้ การวิเคราะห์ในส่วนนี้ได้ศึกษากับงานวิจัยที่มีตัวแปรคุณลักษณะงานวิจัยที่มีค่าเฉลี่ยดัชนีมาตรฐานมากกว่า 2 ค่า ขึ้นไป เนื่องจากในกรณีที่คุณลักษณะใดมีค่าดัชนีมาตรฐานเพียง 1 ค่า ทำให้ไม่มีความแปรปรวนภายในกลุ่ม จึงไม่สามารถทดสอบความแตกต่างของค่าเฉลี่ยรายคู่ได้ ผลการวิเคราะห์ข้อมูลมีรายละเอียดดังต่อไปนี้

3.2.1 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงจำแนกตามตัวแปรคุณลักษณะงานวิจัย

ด้านการพิมพ์และผู้วิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านการพิมพ์และผู้วิจัย ประกอบด้วยคุณลักษณะ 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, ประเภทงานวิจัย และเพศของผู้วิจัย ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัย ได้แก่ ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย และเพศของผู้วิจัย เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = .000-.023$) มีรายละเอียดดังนี้

ตัวแปรประเภทสถาบันที่ผลิตงานวิจัย พบว่า งานวิจัยจากมหาวิทยาลัยราชภัฏ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .823, SD = .106$) รองลงมาคือ สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ($M = .801, SD = .184$), มหาวิทยาลัยในกำกับของรัฐ ($M = .761, SD = .156$) และมหาวิทยาลัยของรัฐ ($M = .731, SD = .238$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยจากมหาวิทยาลัยราชภัฏ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยจากมหาวิทยาลัยในกำกับของรัฐอย่างมีนัยสำคัญทางสถิติ

ตัวแปรสถาบันที่ผลิตงานวิจัย พบว่า งานวิจัยจากมหาวิทยาลัยสุโขทัยธรรมมาธิราชมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .895, SD = .091$) รองลงมาคือ มหาวิทยาลัยนเรศวร ($M = .866, SD = .047$) มหาวิทยาลัยสงขลานครินทร์ ($M = .856, SD = .020$), มหาวิทยาลัยราชภัฏอุบลราชธานี ($M = .849, SD = .111$), มหาวิทยาลัยราชภัฏมหาสารคาม ($M = .845, SD = .092$), มหาวิทยาลัยทักษิณ ($M = .842, SD = .126$), มหาวิทยาลัยราชภัฏกาญจนบุรี ($M = .836, SD = .064$), สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ($M = .801, SD = .184$), มหาวิทยาลัยเกษตรศาสตร์ ($M = .793, SD = .198$), มหาวิทยาลัยเชียงใหม่ ($M = .787, SD = .104$), มหาวิทยาลัยบูรพา ($M = .776, SD = .194$), มหาวิทยาลัยราชภัฏนครราชสีมา ($M = .776, SD = .048$), จุฬาลงกรณ์มหาวิทยาลัย ($M = .757, SD = .142$), มหาวิทยาลัยราชภัฏเพชรบูรณ์ ($M = .755, SD = .116$), มหาวิทยาลัยขอนแก่น ($M = .722, SD = .171$), มหาวิทยาลัยราชภัฏเลย ($M = .713, SD = .221$), มหาวิทยาลัยศรีนครินทรวิโรฒ ($M = .641, SD = .171$), มหาวิทยาลัยราชภัฏเลย ($M = .713, SD = .221$), มหาวิทยาลัยศรีนครินทรวิโรฒ ($M = .641, SD = .171$),

= .146) และมหาวิทยาลัยมหาสารคาม ($M = .591, SD = .257$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยจากมหาวิทยาลัยนเรศวร มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยจากจุฬาลงกรณ์มหาวิทยาลัย อย่างมีนัยสำคัญทางสถิติ รวมถึงงานวิจัยจากมหาวิทยาลัยนเรศวร มหาวิทยาลัยสงขลานครินทร์ และมหาวิทยาลัยสุโขทัยธรรมาราชมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยจากมหาวิทยาลัยมหาสารคาม อย่างมีนัยสำคัญทางสถิติ

ตัวแปรสาขาที่ผลิตงานวิจัย พบว่า งานวิจัยจากกลุ่มสาขาการวิจัยและสถิติ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .825, SD = .120$) รองลงมาคือ รายงานวิจัย ($M = .801, SD = .184$), กลุ่มสาขาการวัดและประเมินผล ($M = .742, SD = .181$) และเทคโนโลยีการวัดทางการศึกษา ($M = .615, SD = .092$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยจากกลุ่มสาขาการวิจัยและสถิติ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยจากกลุ่มสาขาการวัดและประเมินผล อย่างมีนัยสำคัญทางสถิติ

ตัวแปรเพศของผู้วิจัย พบว่า งานวิจัยจากนักวิจัยชาย มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .806, SD = .134$) รองลงมาคือ งานวิจัยจากนักวิจัยหญิง ($M = .778, SD = .164$) และงานวิจัยจากนักวิจัยชายและหญิง ($M = .445, SD = .132$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยจากนักวิจัยชาย และงานวิจัยจากนักวิจัยหญิงมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยจากชายและหญิง อย่างมีนัยสำคัญทางสถิติ

ด้านเนื้อหาสาระงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านเนื้อหาสาระงานวิจัย ประกอบด้วยคุณลักษณะ 18 ตัวแปร ได้แก่ องค์ประกอบที่มุ่งวัด, เนื้อหาที่เกี่ยวข้อง, ภูมิภาค, แหล่งที่มาของตัวอย่าง, ระดับชั้นของตัวอย่าง, การศึกษาแนวคิดการสร้างเครื่องมือวัด, การกำหนดนิยามเชิงปฏิบัติการ, การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, ลักษณะวิธีตอบของเครื่องมือวัด, การตรวจให้คะแนน, จำนวนครั้งที่ใช้ในการทดลอง, จำนวนข้อคำถามในแต่ละฉบับ, จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา, วิธีที่ใช้ในการสร้างเกณฑ์ปกติ และคู่มือการใช้เครื่องมือวัด ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัย ได้แก่ ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, การตรวจให้คะแนน และจำนวนข้อคำถามในแต่ละฉบับ เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = .000-.040$) มีรายละเอียดดังนี้

ตัวแปรภูมิภาค พบว่า งานวิจัยที่มีการศึกษาในภาคใต้ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .853, SD = .127$) รองลงมาคือ งานวิจัยที่มีการศึกษาในภาคเหนือ ($M = .823, SD = .101$), กรุงเทพมหานคร ($M = .810, SD = .162$), ภาคกลาง ($M = .774, SD = .154$), ภาคตะวันออกเฉียงเหนือ ($M = .756, SD = .190$), มากกว่า 1 ภูมิภาค ($M = .754, SD = .130$) และไม่ระบุภูมิภาค ($M = .445, SD = .132$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยที่มีการศึกษาในกรุงเทพมหานคร, ภาคเหนือ และภาคใต้ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่ไม่ระบุภูมิภาคที่ศึกษา อย่างมีนัยสำคัญทางสถิติ

ตัวแปรระดับชั้นของตัวอย่าง พบว่า งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้นมัธยมศึกษาปีที่ 2 มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .865, SD = .067$) รองลงมาคือ ชั้นมัธยมศึกษาตอนต้นและตอนปลาย ($M = .860, SD = .095$), ชั้นมัธยมศึกษาปีที่ 5 ($M = .838, SD =$

.079), ชั้นมัธยมศึกษาปีที่ 1 ($M = .830, SD = .132$), ชั้นมัธยมศึกษาปีที่ 3 ($M = .821, SD = .128$), ชั้นประถมศึกษาและมัธยมศึกษา ($M = .772, SD = .210$), ชั้นมัธยมศึกษาตอนปลาย ($M = .766, SD = .149$), ชั้นมัธยมศึกษาตอนต้น ($M = .732, SD = .154$), ชั้นมัธยมศึกษาปีที่ 6 ($M = .700, SD = .013$) และชั้นมัธยมศึกษาปีที่ 4 ($M = .458, SD = .275$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้นมัธยมศึกษาปีที่ 2, 5 และชั้นมัธยมศึกษาตอนต้น และตอนปลาย มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่ศึกษากับตัวอย่างในชั้นมัธยมศึกษาปีที่ 4 อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้น 1, 2, 3, 5 และตัวอย่างในชั้นมัธยมศึกษาตอนต้นและตอนปลาย มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่ศึกษากับตัวอย่างในชั้นมัธยมศึกษาปีที่ 6 อย่างมีนัยสำคัญทางสถิติ และยังพบว่า งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้น 2, 3 มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่ศึกษากับตัวอย่างในชั้นมัธยมศึกษาตอนต้น อย่างมีนัยสำคัญทางสถิติ

ตัวแปรประเภทของเครื่องมือวัด พบว่า งานวิจัยที่มีการสร้างเครื่องมือแบบสังเกต มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .889, SD = .039$) รองลงมาคือ มาตรฐานค่า ($M = .838, SD = .116$), แบบสอบอัตนัย ($M = .834, SD = .104$), อื่น ๆ ($M = .825, SD = .132$), แบบสอบหลายตัวเลือก ($M = .785, SD = .151$), แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า ($M = .753, SD = .113$), แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า ($M = .712, SD = .177$), แบบสอบภาคปฏิบัติ ($M = .705, SD = .248$), แบบสอบอัตนัยประยุกต์ (แบบสอบอัตนัยประยุกต์ (MEQ)) ($M = .671, SD = .305$) และแบบตรวจสอบรายการ ($M = .662, SD = .089$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่มีการสร้างเครื่องมือแบบสังเกต มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่สร้างเครื่องมือแบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยที่มีการสร้างเครื่องมือแบบสอบอัตนัย มาตรฐานค่า และแบบสังเกตมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่สร้างเครื่องมือแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า อย่างมีนัยสำคัญทางสถิติ

ตัวแปรจำนวนตัวเลือก/เกณฑ์ พบว่า งานวิจัยที่มีการสร้างเครื่องมือที่มีมากกว่า 5 ตัวเลือก/เกณฑ์ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .939, SD = .030$) รองลงมาคือ ใช้มากกว่า 1 ประเภท ($M = .835, SD = .139$), ไม่มีตัวเลือก/เกณฑ์ ($M = .831, SD = .131$), 5 ตัวเลือก/เกณฑ์ ($M = .803, SD = .117$), 3 ตัวเลือก/เกณฑ์ ($M = .734, SD = .169$), 4 ตัวเลือก/เกณฑ์ ($M = .726, SD = .196$) และ 2 ตัวเลือก/เกณฑ์ ($M = .662, SD = .089$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่มีการสร้างเครื่องมือที่มีมากกว่า 5 ตัวเลือก/เกณฑ์ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่มีการสร้างเครื่องมือ 3 ตัวเลือก/เกณฑ์ อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยที่มีการสร้างเครื่องมือที่ไม่มีตัวเลือก/เกณฑ์ และงานวิจัยที่มีการสร้างเครื่องมือที่มีมากกว่า 5 ตัวเลือก/เกณฑ์ มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่มีการสร้างเครื่องมือ 4 ตัวเลือก/เกณฑ์ อย่างมีนัยสำคัญทางสถิติ

ตัวแปรการตรวจให้คะแนน พบว่า งานวิจัยที่มีการตรวจให้คะแนนแบบสองค่า และมากกว่าสองค่า มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงที่สุด ($M = .923, SD = .068$) รองลงมาคือ ให้คะแนนมากกว่าสองค่า ($M = .807, SD = .125$) และให้คะแนนแบบสองค่า ($M = .685, SD = .213$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่มีการตรวจให้คะแนนแบบสองค่า มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่มีการตรวจให้คะแนนแบบสองค่าอย่างมีนัยสำคัญ

ทางสถิติ นอกจากนี้ งานวิจัยที่มีการตรวจให้คะแนนแบบสองค่า และมากกว่าสองค่า มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่มีการตรวจให้คะแนนแบบสองค่า และงานวิจัยที่มีการตรวจให้คะแนนแบบมากกว่าสองค่า อย่างมีนัยสำคัญทางสถิติ

ตัวแปรจำนวนข้อคำถามในแต่ละฉบับ พบว่า งานวิจัยที่มีจำนวนข้อคำถามในแต่ละฉบับมากกว่า 30 ข้อ ขึ้นไป มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงสุด ($M = .845, SD = .098$) รองลงมาคือ น้อยกว่าหรือเท่ากับ 30 ข้อ ($M = .771, SD = .168$) ผลการเปรียบเทียบค่าเฉลี่ยดัชนีมาตรฐานความเที่ยง พบว่า งานวิจัยที่มีจำนวนข้อคำถามในแต่ละฉบับมากกว่า 30 ข้อ ขึ้นไป มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่า งานวิจัยที่มีจำนวนข้อคำถามในแต่ละฉบับน้อยกว่าหรือเท่ากับ 30 ข้อ อย่างมีนัยสำคัญทางสถิติ

ด้านวิธีวิทยาการวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านวิธีวิทยาการวิจัย ประกอบด้วยคุณลักษณะ 2 ตัวแปร ได้แก่ วิธีการเลือกตัวอย่าง และจำนวนผู้สอบในแต่ละฉบับ ผลการวิเคราะห์พบว่า ตัวแปรวิธีการเลือกตัวอย่าง เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p=.002$) กล่าวคือ เมื่อพิจารณาข้อมูลพื้นฐาน พบว่า งานวิจัยที่ใช้ประชากร/กลุ่มเป้าหมายในการศึกษา ($M = .935, SD = .007$) มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงสุด รองลงมาคือ สุ่มอย่างง่าย ($M = .864, SD = .048$), สุ่มแบบหลายขั้นตอน ($M = .831, SD = .110$), สุ่มแบบสองขั้นตอน ($M = .790, SD = .147$), ไม่ระบุ ($M = .755, SD = .147$), เลือกอย่างเจาะจง ($M = .754, SD = .169$), สุ่มแบบกลุ่ม ($M = .745, SD = .106$) และสุ่มแบบแบ่งชั้น ($M = .676, SD = .272$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่ใช้ประชากร/กลุ่มเป้าหมาย ค่าเฉลี่ยมีดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่มีการสุ่มแบบแบ่งชั้น, สุ่มแบบสองขั้นตอน, สุ่มแบบหลายขั้นตอน, เลือกอย่างเจาะจง และไม่ระบุวิธีการเลือกตัวอย่าง อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยที่ใช้วิธีการสุ่มอย่างง่าย และสุ่มแบบสองขั้นตอน มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่างานวิจัยที่ไม่ระบุวิธีการเลือกตัวอย่าง อย่างมีนัยสำคัญทางสถิติ

ด้านคุณภาพงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านคุณภาพงานวิจัย มีตัวแปรคุณลักษณะ 1 ตัวแปร คือ ระดับคุณภาพงานวิจัย ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัยด้านคุณภาพงานวิจัยไม่มีผลต่อความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยง กล่าวคือ งานวิจัยที่มีระดับคุณภาพต่างกันมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงไม่แตกต่างกัน รายละเอียดดังตารางที่ 4.7

ตารางที่ 4.7 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงจำแนกตามตัวแปร
คุณลักษณะงานวิจัย

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p	
ด้านการพิมพ์และผู้วิจัย									
ปีที่พิมพ์ งานวิจัย	(1) พ.ศ. 2546-2550	54	.781	.158	3.503	.032	2.822	.062	
	(2) พ.ศ. 2551-2555	60	.738	.201					
	(3) พ.ศ. 2556-2561	102	.801	.136					
ประเภท สถาบันที่ผลิต งานวิจัย	(1) มหาวิทยาลัยของรัฐ	35	.731	.238	8.841	.000	3.257	.023	
	(2) มหาวิทยาลัยในกำกับ ของรัฐ	100	.761	.156					(3) > (2)
	(3) มหาวิทยาลัยราชภัฏ	69	.823	.106					
	(4) สทศ.	12	.801	.184					
สถาบันที่ผลิต งานวิจัย	(1) จุฬาลงกรณ์มหาวิทยาลัย	45	.757	.142	4.863	.000	3.539	.000	
	(2) ม.เกษตรศาสตร์	13	.793	.198					(6) > (1)
	(3) ม.เชียงใหม่	7	.787	.104					(6), (10) และ (11) > (8)
	(4) ม.ขอนแก่น	12	.722	.171					
	(5) ม.ทักษิณ	13	.842	.126					
	(6) ม.นเรศวร	10	.866	.047					
	(7) ม.บูรพา	4	.776	.194					
	(8) ม.มหาสารคาม	18	.591	.257					
	(9) ม.ศรีนครินทรวิโรฒ	8	.641	.146					
	(10) ม.สงขลานครินทร์	2	.856	.020					
	(11) ม.สุโขทัยธรรมมาธิราช	8	.895	.091					
	(12) มรภ.กาญจนบุรี	7	.836	.064					
	(13) มรภ.นครราชสีมา	6	.776	.048					
	(14) มรภ.เพชรบูรณ์	7	.755	.116					
	(15) มรภ.มหาสารคาม	21	.845	.092					
	(16) มรภ.เลย	3	.713	.221					
	(17) มรภ.อุบลราชธานี	20	.849	.111					
	(18) สทศ.	12	.801	.184					
สาขาที่ผลิต งานวิจัย	(1) กลุ่มสาขาการวัดและ ประเมินผล	113	.742	.181	5.271	.002	5.306	.002	
	(2) กลุ่มสาขาการวิจัยและ สถิติ	89	.825	.120					(2) > (1)
	(3) เทคโนโลยีการวัดทาง การศึกษา	2	.615	.092					
	(4) รายงานวิจัย	12	.801	.184					
ประเภท งานวิจัย	(1) วิทยานิพนธ์ ปริญญาโท	176	.780	.167	.219	.803	.375	.688	
	(2) วิทยานิพนธ์ ปริญญาเอก	28	.756	.132					
	(3) รายงานวิจัย	12	.801	.184					

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
เพศของผู้วิจัย	(1) ชาย	27	.806	.134	.238	.788	4.739	.010
	(2) หญิง	187	.778	.164				
	(3) ชายและหญิง	2	.445	.132				
ด้านเนื้อหาสาระงานวิจัย								
องค์ประกอบ ที่มุ่งวัด	(1) จิตสำนึกต่อโลก	3	.877	.015	2.560	.002	1.162	.307
	(2) ความรู้เรื่องพลเมือง	2	.965	.035				
	(3) การคิดอย่างมี วิจารณ์ญาณ	13	.749	.148				
	(4) การแก้ปัญหา	37	.735	.240				
	(5) การสร้างสรรค์และ นวัตกรรม	50	.771	.153				
	(6) การสื่อสาร	32	.813	.128				
	(7) ทักษะด้านการทำงานเป็น ทีม ความร่วมมือ	9	.850	.086				
	(8) ภาวะผู้นำ	4	.832	.070				
	(9) ความรับผิดชอบ	26	.775	.146				
	(10) การรู้เท่าทันสื่อและ สารสนเทศ	7	.830	.119				
	(11) การรู้เท่าทันเทคโนโลยี	11	.842	.127				
	(12) การคิดอย่างมี วิจารณ์ญาณและการ แก้ปัญหา	9	.755	.129				
	(13) การแก้ปัญหาย่าง สร้างสรรค์	2	.730	.198				
	(14) การสื่อสารและการ ร่วมมือ	8	.708	.190				
	(15) ภาวะผู้นำและความ รับผิดชอบ	3	.667	.080				
เนื้อหาที่ เกี่ยวข้อง	(1) คณิตศาสตร์	45	.769	.227	3.768	.001	1.762	.108
	(2) วิทยาศาสตร์	26	.819	.109				
	(3) สังคมศึกษา ศาสนา และวัฒนธรรม	1*	.760	-				
	(4) ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์)	2	.975	.021				
	(5) การงานอาชีพและ เทคโนโลยี	3	.941	.004				
	(6) ภาษาต่างประเทศ	11	.802	.130				
	(7) ไม่สังกัดกลุ่มสาระ (content free)	126	.762	.147				
	(8) มากกว่า 1 สาระ	2	.910	.042				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p	
ภูมิภาค	(1) กทม.	21	.810	.162	1.836	.094	3.005	.008	
	(2) ภาคเหนือ	34	.823	.101					(1), (2) และ (5) > (7)
	(3) ภาคตะวันออกเฉียงเหนือ	81	.756	.190					
	(4) ภาคกลาง	35	.774	.154					
	(5) ภาคใต้	15	.853	.127					
	(6) มากกว่า 1 ภูมิภาค	28	.754	.130					
	(7) ไม่ระบุ	2	.445	.132					
แหล่งที่มาของ ตัวอย่าง	(1) สำนักงานคณะกรรมการ การศึกษาขั้นพื้นฐาน (สพฐ.)	172	.781	.165	1.168	.325	1.003	.424	
	(2) สำนักงานคณะกรรมการ ส่งเสริมการศึกษาเอกชน	3	.817	.145					
	(3) สังกัดองค์กรปกครอง ส่วนท้องถิ่น	4	.780	.065					
	(4) สังกัดคณะกรรมการการ อุดมศึกษา (สกอ.)	4	.625	.264					
	(5) อื่น ๆ	2	.935	.007					
	(6) มากกว่า 1 สังกัด	7	.739	.143					
	(7) ไม่ระบุ	24	.783	.150					
ระดับชั้นของ ตัวอย่าง	(1) ชั้นมัธยมศึกษาปีที่ 1	21	.830	.132	5.716	.000	8.785	.000	
	(2) ชั้นมัธยมศึกษาปีที่ 2	17	.865	.067					(2), (5) และ (9) > (4)
	(3) ชั้นมัธยมศึกษาปีที่ 3	43	.821	.128					(1), (2), (3), (5) และ (9) > (6)
	(4) ชั้นมัธยมศึกษาปีที่ 4	10	.458	.275					(2) และ (3) > (7)
	(5) ชั้นมัธยมศึกษาปีที่ 5	25	.838	.079					
	(6) ชั้นมัธยมศึกษาปีที่ 6	4	.700	.013					
	(7) ชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน)	58	.732	.154					
	(8) ชั้นมัธยมศึกษาตอน ปลาย (มากกว่า 1 ชั้นเรียน)	22	.766	.149					
	(9) ชั้นมัธยมศึกษาตอนต้น และตอนปลาย	8	.860	.095					
	(10) ชั้นประถมศึกษาและ มัธยมศึกษา	8	.772	.210					
การศึกษา แนวคิดการ สร้างเครื่องมือ วัด	(1) มี เฉพาะการวิเคราะห์ คุณภาพเครื่องมือ	14	.818	.129	.712	.492	.448	.639	
	(2) มี เฉพาะวิธีการสร้าง เครื่องมือ	54	.778	.170					
	(3) มี ทั้งวิธีการสร้างและ การวิเคราะห์คุณภาพ	148	.775	.164					

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p	
การสร้าง เครื่องมือวัด	(1) ผู้วิจัยพัฒนาเครื่องมือวัด ขึ้นเอง	214	.779	.163	.052	.821	.355	.552	
	(2) ผู้วิจัยปรับปรุงเครื่องมือ วัดจากงานวิจัยต่างประเทศ	2	.710	.156					
การกำหนด นิยามเชิง ปฏิบัติการ	(1) ไม่มี	7	.769	.136	2.034	.133	.642	.527	
	(2) มี แต่ไม่นำเสนอ	20	.818	.116					
	(3) มี และนำเสนอ	189	.775	.168					
การสร้าง ตารางกำหนด โครงสร้างของ เครื่องมือวัด	(1) ไม่มี	41	.804	.146	1.089	.338	.686	.505	
	(2) มี แต่ไม่นำเสนอ	16	.787	.122					
	(3) มี และนำเสนอ	159	.771	.171					
ประเภทของ เครื่องมือวัด	(1) แบบสอบหลายตัวเลือก	19	.785	.151	11.810	.000	3.684	.000	
	(2) แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า	26	.712	.177					(8) > (2)
	(3) แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า	46	.753	.113					(4), (7) และ (8) > (3)
	(4) แบบสอบอัตนัย	45	.834	.104					
	(5) แบบสอบภาคปฏิบัติ	7	.705	.248					
	(6) แบบตรวจสอบรายการ	3	.662	.089					
	(7) มาตรฐานประมาณค่า	39	.838	.116					
	(8) แบบสังเกต	5	.889	.039					
	(9) แบบสอบอัตนัยประยุกต์ (MEQ)	19	.671	.305					
	(10) อื่น ๆ	7	.825	.132					
จำนวน ตัวเลือก/ เกณฑ์	(1) ไม่มีตัวเลือก/เกณฑ์	65	.831	.131	2.535	.022	3.746	.001	
	(2) 2 ตัวเลือก/เกณฑ์	3	.662	.089					(6) > (3)
	(3) 3 ตัวเลือก/เกณฑ์	22	.734	.169					(1) และ (6) > (4)
	(4) 4 ตัวเลือก/เกณฑ์	73	.726	.196					
	(5) 5 ตัวเลือก/เกณฑ์	47	.803	.117					
	(6) มากกว่า 5 ตัวเลือก/เกณฑ์	2	.939	.030					
	(7) ใช้มากกว่า 1 ประเภท	4	.835	.139					
ลักษณะวิธี ตอบของ เครื่องมือวัด	(1) แบบเขียนตอบ	195	.777	.162	2.652	.034	1.492	.206	
	(2) คอมพิวเตอร์	9	.836	.117					
	(3) ปฏิบัติจริง	7	.741	.231					
	(4) ปากเปล่า	2	.975	.021					
	(5) อื่น ๆ	3	.662	.089					
การตรวจให้ คะแนน	(1) ให้คะแนนแบบสองค่า	57	.685	.213	12.839	.000	16.566	.000	
	(2) ให้คะแนนมากกว่าสองค่า	152	.807	.125					2) > (1)
	(3) ให้คะแนนแบบสองค่า และมากกว่าสองค่า	7	.923	.068					(3) > (1) และ (2)

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p	
จำนวนครั้งที่ ใช้ในการ ทดลอง	(1) 1 ครั้ง	10	.856	.072	1.902	.111	1.956	.102	
	(2) 2 ครั้ง	96	.784	.148					
	(3) 3 ครั้ง	93	.765	.178					
	(4) 4 ครั้ง	11	.838	.140					
	(5) 5 ครั้ง	6	.660	.229					
จำนวนข้อ คำถามในแต่ละ ฉบับ	(1) น้อยกว่าหรือเท่ากับ 30 ข้อ	193	.771	.168	5.474	.020	4.290	.040	
	(2) มากกว่า 30 ข้อ ขึ้นไป	23	.845	.098					(2) > (1)
จำนวน ผู้เชี่ยวชาญใน การตรวจสอบ ความตรงตาม เนื้อหา	(1) น้อยกว่าหรือเท่ากับ 5 คน	144	.784	.178	3.226	.074	.482	.488	
	(2) มากกว่า 5 คน ขึ้นไป	72	.768	.129					
วิธีที่ใช้ในการ สร้างเกณฑ์ ปกติ	(1) วิธีเปอร์เซ็นต์ไทล์	6	.728	.261	4.000	.004	1.210	.308	
	(2) วิธีคะแนนมาตรฐานที่	1*	.914	-					
	(3) วิธีคะแนนมาตรฐานที่	121	.781	.124					
	(4) วิธีเปอร์เซ็นต์ไทล์และ	3	.590	.106					
	คะแนนมาตรฐานที่รวมกัน								
	(5) วิธีเปอร์เซ็นต์ไทล์และสเต	5	.811	.184					
โนนรรวมกัน									
คู่มือการใช้ เครื่องมือวัด	(1) มี	43	.831	.117	5.649	.018	5.593	.019	
	(2) ไม่มี	173	.766	.170					
ด้านวิธีวิทยาการวิจัย									
วิธีการเลือก ตัวอย่าง	(1) ใช้ประชากร/กลุ่มเป้าหมาย	2	.935	.007	9.223	.000	3.352	.002	
	(2) สุ่มอย่างง่าย	5	.864	.048					(1) > (4), (5), (6), (7) และ (8)
	(3) สุ่มแบบกลุ่ม	2	.745	.106					(2) และ (5) > (8)
	(4) สุ่มแบบแบ่งชั้น	25	.676	.272					
	(5) สุ่มแบบสองขั้นตอน	36	.790	.147					
	(6) สุ่มแบบหลายขั้นตอน	65	.831	.110					
	(7) เลือกอย่างเจาะจง	16	.754	.169					
	(8) ไม่ระบุ	65	.755	.147					
จำนวนผู้สอบ ในแต่ละฉบับ	(1) น้อยกว่าหรือเท่ากับ 200 คน	72	.759	.197	4.649	.032	1.476	.226	
	(2) มากกว่า 200 คน ขึ้นไป	144	.788	.143					
ด้านคุณภาพงานวิจัย									
ระดับคุณภาพ งานวิจัย	(1) ดี	175	.785	.167	.007	.935	1.671	.198	
	(2) ดีมาก	41	.749	.144					

* ไม่ได้นำเข้าร่วมวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA)

3.2.2 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรงจำแนกตามตัวแปรคุณลักษณะงานวิจัย

ด้านการพิมพ์และผู้วิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านการพิมพ์และผู้วิจัย ประกอบด้วยคุณลักษณะ 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, ประเภทงานวิจัย และเพศของผู้วิจัย ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัยด้านการพิมพ์และผู้วิจัยไม่มีผลต่อความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรง กล่าวคือ งานวิจัยที่มีปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, ประเภทงานวิจัย และเพศของผู้วิจัยที่แตกต่างกัน มีค่าเฉลี่ยดัชนีมาตรฐานความตรงไม่แตกต่างกัน

ด้านเนื้อหาสาระงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านเนื้อหาสาระงานวิจัย ประกอบด้วยคุณลักษณะ 18 ตัวแปร ได้แก่ องค์กรประกอบที่มุ่งวัด, เนื้อหาที่เกี่ยวข้อง, ภูมิภาค, แหล่งที่มาของตัวอย่าง, ระดับชั้นของตัวอย่าง, การศึกษาแนวคิดการสร้างเครื่องมือวัด, การกำหนดนิยามเชิงปฏิบัติการ, การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, ลักษณะวิธีตอบของเครื่องมือวัด, การตรวจให้คะแนน, จำนวนครั้งที่ใช้ในการทดลอง, จำนวนข้อคำถามในแต่ละฉบับ, จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา, วิธีที่ใช้ในการสร้างเกณฑ์ปกติ และคู่มือการใช้เครื่องมือวัด ผลการวิเคราะห์พบว่า มีเพียงตัวแปรคุณลักษณะงานวิจัย 1 ตัวแปร คือ เนื้อหาที่เกี่ยวข้อง เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานความตรงแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p=.018$) มีรายละเอียดดังนี้

ตัวแปรเนื้อหาที่เกี่ยวข้อง พบว่า งานวิจัยที่มีการสร้างเครื่องมือวัดที่มีเนื้อหาเกี่ยวข้องกับวิทยาศาสตร์ มีค่าเฉลี่ยดัชนีมาตรฐานความตรงสูงสุด ($M = .995, SD = .000$) รองลงมาคือ สังคมศึกษา ศาสนา และวัฒนธรรม ($M = .720, SD = -$), ไม่สังกัดกลุ่มสาระ (content free) ($M = .600, SD = .164$), คณิตศาสตร์ ($M = .540, SD = .186$) และภาษาต่างประเทศ ($M = .533, SD = .229$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยที่มีการสร้างเครื่องมือวัดที่มีเนื้อหาเกี่ยวข้องกับวิทยาศาสตร์ มีค่าเฉลี่ยดัชนีมาตรฐานความตรงสูงกว่างานวิจัยที่มีการสร้างเครื่องมือวัดที่มีเนื้อหาเกี่ยวข้องกับคณิตศาสตร์, ภาษาต่างประเทศ และไม่สังกัดกลุ่มสาระ (content free) อย่างมีนัยสำคัญทางสถิติ

ด้านวิธีวิทยาการวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านวิธีวิทยาการวิจัย ประกอบด้วยคุณลักษณะ 2 ตัวแปร ได้แก่ วิธีการเลือกตัวอย่างและจำนวนผู้สอบในแต่ละฉบับ ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัยด้านวิธีวิทยาการวิจัย ไม่มีผลต่อความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรง กล่าวคือ วิธีการเลือกตัวอย่าง และจำนวนผู้สอบในแต่ละฉบับที่แตกต่างกัน มีค่าเฉลี่ยดัชนีมาตรฐานความตรงไม่แตกต่างกัน

ด้านคุณภาพงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรงจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านคุณภาพงานวิจัย มีตัวแปรคุณลักษณะ 1 ตัวแปร คือ ระดับคุณภาพงานวิจัย แต่ทั้งนี้ไม่มีการทดสอบเนื่องจากค่าเฉลี่ยดัชนีมาตรฐานความตรงของคุณภาพงานวิจัยในระดับดีมาก มีเพียง 1

ค่า ทำให้ไม่มีความแปรปรวนภายในกลุ่ม จึงไม่สามารถทดสอบความแตกต่างของค่าเฉลี่ยได้ รายละเอียดดังตารางที่ 4.8

ตารางที่ 4.8 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความตรงจำแนกตามตัวแปรคุณลักษณะงานวิจัย

คุณลักษณะงานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
ด้านการพิมพ์และผู้วิจัย								
ปีที่พิมพ์งานวิจัย	(1) พ.ศ. 2546-2550	10	.619	.148	3.558	.039	1.124	.336
	(2) พ.ศ. 2551-2555	21	.555	.189				
	(3) พ.ศ. 2556-2561	8	.674	.268				
ประเภทสถาบันที่ผลิตงานวิจัย	(1) มหาวิทยาลัยของรัฐ	13	.581	.180	3.976	.015	1.282	.296
	(2) มหาวิทยาลัยในกำกับของรัฐ	12	.584	.141				
	(3) มหาวิทยาลัยราชภัฏ	12	.660	.260				
	(4) สทศ.	2	.379	.016				
สถาบันที่ผลิตงานวิจัย	(1) ม.เกษตรศาสตร์	5	.686	.125	2.449	0.044	1.646	.165
	(2) ม.เชียงใหม่	2	.511	.011				
	(3) ม.ขอนแก่น	2	.622	.139				
	(4) ม.ทักษิณ	5	.548	.274				
	(5) ม.นเรศวร	1*	.657	-				
	(6) ม.บูรพา	1*	.461	-				
	(7) ม.มหาสารคาม	12	.575	.186				
	(8) ม.ศรีนครินทรวิโรฒ	1*	.364	-				
	(9) มรภ.กาญจนบุรี	4	.815	.221				
	(10) มรภ.เลย	3	.493	.068				
	(11) มรภ.อุบลราชธานี	1*	.930	-				
	(12) สทศ.	2	.379	.016				
สาขาที่ผลิตงานวิจัย	(1) กลุ่มสาขาการวัดและประเมินผล	23	.562	.187	1.924	.161	3.220	.052
	(2) กลุ่มสาขาการวิจัยและสถิติ	14	.683	.197				
	(3) รายงานวิจัย	2	.379	.016				
ประเภทงานวิจัย	(1) วิทยานิพนธ์ ปริญญาโท	37	.608	.197	3.489	.070	2.631	.113
	(2) รายงานวิจัย	2	.379	.016				
เพศของผู้วิจัย	(1) ชาย	2	.637	.386	2.154	.151	.088	.768
	(2) หญิง	37	.594	.193				
ด้านเนื้อหาสาระงานวิจัย								
องค์ประกอบที่มุ่งวัด	(1) การคิดอย่างมีวิจารณญาณ	3	.633	.101	1.231	.320	1.743	.124
	(2) การแก้ปัญหา	12	.544	.170				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p	
	(3) การสร้างสรรค์และ นวัตกรรม	4	.528	.088					
	(4) การสื่อสาร	6	.566	.246					
	(5) ทักษะด้านการทำงานเป็น ทีม ความร่วมมือ	1*	.995	-					
	(6) ภาวะผู้นำ	1*	.543	-					
	(7) ความรับผิดชอบ	8	.704	.210					
	(8) การรู้เท่าทันเทคโนโลยี	1*	.461	-					
	(9) การคิดอย่างมี วิจารณญาณและการ แก้ปัญหา	2	.379	.016					
	(10) การแก้ปัญหาอย่าง สร้างสรรค์	1*	.910	-					
เนื้อหาที่ เกี่ยวข้อง	(1) คณิตศาสตร์	10	.540	.186	1.279	.297	3.865	.018	
	(2) วิทยาศาสตร์	2	.995	.000					(2) > (1), (4) และ (5)
	(3) สังคมศึกษา ศาสนา และวัฒนธรรม*	1*	.720	-					
	(4) ภาษาต่างประเทศ	7	.533	.229					
	(5) ไม่สังกัดกลุ่มสาระ (content free)	19	.600	.164					
ภูมิภาค	(1) กทม.	6	.546	.130	2.280	.081	.801	.533	
	(2) ภาคเหนือ	3	.559	.085					
	(3) ภาคตะวันออกเฉียงเหนือ	18	.586	.181					
	(4) ภาคกลาง	7	.714	.262					
	(5) ภาคใต้	5	.548	.274					
แหล่งที่มาของ ตัวอย่าง	(1) สำนักงานคณะกรรมการ การศึกษาขั้นพื้นฐาน (สพฐ.)	31	.611	.214	8.002	.001	1.332	.277	
	(2) สังกัดคณะกรรมการการ อุดมศึกษา (สกอ.)	4	.630	.000					
	(3) ไม่ระบุ	4	.445	.077					
ระดับชั้นของ ตัวอย่าง	(1) ชั้นมัธยมศึกษาปีที่ 1	6	.733	.267	4.746	.002	1.870	.117	
	(2) ชั้นมัธยมศึกษาปีที่ 2	1*	.720	-					
	(3) ชั้นมัธยมศึกษาปีที่ 3	7	.549	.243					
	(4) ชั้นมัธยมศึกษาปีที่ 4	9	.497	.133					
	(5) ชั้นมัธยมศึกษาปีที่ 6	4	.445	.077					
	(6) ชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน)	10	.666	.124					
	(7) ชั้นมัธยมศึกษาตอน ปลาย (มากกว่า 1 ชั้นเรียน)	2	.686	.317					

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
การศึกษา แนวคิดการ สร้างเครื่องมือ วัด	(1) มี เฉพาะการวิเคราะห์ คุณภาพเครื่องมือ (2) มี เฉพาะวิธีการสร้าง เครื่องมือ (3) มี ทั้งวิธีการสร้างและ การวิเคราะห์คุณภาพ	2 11 26	.860 .579 .582	.085 .086 .226	6.019	.006	1.958	.156
การสร้าง เครื่องมือวัด	(1) ผู้วิจัยพัฒนาเครื่องมือวัด ขึ้นเอง (2) ผู้วิจัยปรับปรุงเครื่องมือ วัดจากงานวิจัยต่างประเทศ	39 -	.596 -	.199 -	ไม่ทดสอบ			
การกำหนด นิยามเชิง ปฏิบัติการ	(1) ไม่มี (2) มี แต่ไม่นำเสนอ (3) มี และนำเสนอ	2 8 29	.379 .512 .634	.016 .220 .186	1.668	.203	2.663	.083
การสร้าง ตารางกำหนด โครงสร้างของ เครื่องมือวัด	(1) ไม่มี (2) มี แต่ไม่นำเสนอ (3) มี และนำเสนอ	3 7 29	.838 .713 .543	.273 .204 .164	.938	.401	5.542	.008
ประเภทของ เครื่องมือวัด	(1) แบบสอบหลายตัวเลือก (2) แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า (3) แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า (4) แบบสอบอัตนัย (5) มาตรฐานค่า (6) แบบสอบอัตนัยประยุกต์ (MEQ)	4 5 6 7 7 10	.484 .704 .583 .533 .736 .540	.134 .122 .198 . .209 .186	.527	.754	1.689	.165
ลักษณะวิธี ตอบของ เครื่องมือวัด	(1) แบบเขียนตอบ (2) คอมพิวเตอร์ (3) ปฏิบัติจริง (4) ปากเปล่า (5) อื่น ๆ	39 - - - -	.596 - - - -	.199 - - - -	ไม่ทดสอบ			
จำนวน ตัวเลือก/เกณฑ์	(1) ไม่มีตัวเลือก/เกณฑ์ (2) 3 ตัวเลือก/เกณฑ์ (3) 4 ตัวเลือก/เกณฑ์ (4) 5 ตัวเลือก/เกณฑ์ (5) ใช้มากกว่า 1 ประเภท	8 5 17 8 1*	.583 .637 .579 .629 .523	.254 .207 .152 .262 -	2.274	.098	.179	.910
การตรวจให้ คะแนน	(1) ให้คะแนนแบบสองค่า (2) ให้คะแนนมากกว่าสอง ค่า	18 21	.552 .634	.156 .226	4.030	.052	1.696	.201

คุณลักษณะงานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
จำนวนครั้งที่ใช้ในการทดลอง	(1) 1 ครั้ง	2	.379	.016	3.127	.027	1.336	.277
	(2) 2 ครั้ง	5	.676	.182				
	(3) 3 ครั้ง	23	.594	.223				
	(4) 4 ครั้ง	4	.501	.058				
	(5) 5 ครั้ง	5	.686	.125				
จำนวนข้อคำถามในแต่ละฉบับ	(1) น้อยกว่าหรือเท่ากับ 30 ข้อ	33	.598	.207	.407	.527	.016	.901
	(2) มากกว่า 30 ข้อ ขึ้นไป	6	.586	.163				
จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา	(1) น้อยกว่าหรือเท่ากับ 5 คน	31	.620	.189	.019	.892	2.245	.143
	(2) มากกว่า 5 คน ขึ้นไป	8	.504	.221				
วิธีที่ใช้ในการสร้างเกณฑ์ปกติ	(1) วิธีเปอร์เซ็นต์ไทล์	4	.630	.000	4.566	.017	.601	.618
	(2) วิธีคะแนนมาตรฐานที่	1*	.461	-				
	(3) วิธีคะแนนมาตรฐานที่ปกติ	20	.628	.223				
	(4) ไม่มีการสร้างเกณฑ์ปกติ	14	.550	.192				
คู่มือการใช้เครื่องมือวัด	(1) มี	3	.374	.014	5.152	.029	4.430	.042
	(2) ไม่มี	36	.614	.196				
ด้านวิธีวิทยาการวิจัย								
วิธีการเลือกตัวอย่าง	(1) สุ่มแบบแบ่งชั้น	12	.582	.224	.325	.807	.692	.563
	(2) สุ่มแบบสองขั้นตอน	9	.535	.215				
	(3) สุ่มแบบหลายขั้นตอน	9	.602	.182				
	(4) ไม่ระบุ	9	.669	.169				
จำนวนผู้สอบในแต่ละฉบับ	(1) น้อยกว่าหรือเท่ากับ 200 คน	24	.603	.183	2.077	.158	.079	.780
	(2) มากกว่า 200 คน ขึ้นไป	15	.584	.227				
ด้านคุณภาพงานวิจัย								
ระดับคุณภาพงานวิจัย	(1) ดี	38	.599	.200	ไม่ทดสอบ			
	(2) ดีมาก	1*	.461	-				

* ไม่ได้นำเข้าร่วมวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA)

3.2.3 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยากจำแนกตามตัวแปรคุณลักษณะงานวิจัย

ด้านการพิมพ์และผู้วิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยากจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านการพิมพ์และผู้วิจัย ประกอบด้วยคุณลักษณะ 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, ประเภทงานวิจัย และเพศของผู้วิจัย ผลการวิเคราะห์พบว่า มีเพียงตัวแปร 2 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย และประเภทงานวิจัย เป็น

ตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานความยากแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p=.011-.032$) มีรายละเอียดดังนี้

ตัวแปรปีที่พิมพ์งานวิจัย พบว่า งานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2556-2561 มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงสุด ($M = .566, SD = .117$) รองลงมาคือ งานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2546-2550 ($M = .529, SD = .115$) และ พ.ศ. 2551-2555 ($M = .499, SD = .097$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2556-2561 มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงกว่างานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2551-2555 อย่างมีนัยสำคัญทางสถิติ

ตัวแปรประเภทงานวิจัย พบว่า งานวิจัยที่เป็นวิทยานิพนธ์ปริญญาเอก ค่าเฉลี่ยดัชนีมาตรฐานความยากสูงสุด ($M = .616, SD = .084$) รองลงมาคือ วิทยานิพนธ์ปริญญาโท ($M = .526, SD = .115$) และรายงานวิจัย ($M = .479, SD = .066$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยที่เป็นวิทยานิพนธ์ปริญญาเอก มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงกว่างานวิจัยที่เป็นวิทยานิพนธ์ปริญญาโท อย่างมีนัยสำคัญทางสถิติ

ด้านเนื้อหาสาระงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยากจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านเนื้อหาสาระงานวิจัย ประกอบด้วยคุณลักษณะ 18 ตัวแปร ได้แก่ องค์ประกอบที่มุ่งวัด, เนื้อหาที่เกี่ยวข้อง, ภูมิภาค, แหล่งที่มาของตัวอย่าง, ระดับชั้นของตัวอย่าง, การศึกษาแนวความคิด, สร้างเครื่องมือวัด, การกำหนดนิยามเชิงปฏิบัติการ, การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, ลักษณะวิธีตอบของเครื่องมือวัด, การตรวจให้คะแนน, จำนวนครั้งที่ใช้ในการทดลอง, จำนวนข้อคำถามในแต่ละฉบับ, จำนวนผู้เชี่ยวชาญในการตรวจสอบความยากตามเนื้อหา, วิธีที่ใช้ในการสร้างเกณฑ์ปกติ และคู่มือการใช้เครื่องมือวัด ผลการวิเคราะห์พบว่า มีเพียงตัวแปรคุณลักษณะงานวิจัย ได้แก่ เนื้อหาที่เกี่ยวข้อง, การสร้างเครื่องมือวัด, ประเภทของเครื่องมือวัด และจำนวนตัวเลือก/เกณฑ์ เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานความยากแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p=.001-.039$) มีรายละเอียดดังนี้

ตัวแปรเนื้อหาที่เกี่ยวข้อง พบว่า งานวิจัยที่มีการสร้างเครื่องมือวัดที่มีเนื้อหาเกี่ยวข้องกับงานอาชีพและเทคโนโลยี มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงสุด ($M = .601, SD = -$) รองลงมาคือ ไม่สังกัดกลุ่มสาระ (content free) ($M = .560, SD = .118$), คณิตศาสตร์ ($M = .557, SD = .088$), วิทยาศาสตร์ ($M = .453, SD = .118$), ภาษาต่างประเทศ ($M = .446, SD = .044$), ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์) ($M = .376, SD = .055$) และสังคมศึกษา ศาสนา และวัฒนธรรม ($M = .352, SD = -$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni ไม่พบคู่ที่มีค่าดัชนีมาตรฐานความยากแตกต่างกัน

ตัวแปรการสร้างเครื่องมือวัด พบว่า งานวิจัยที่มีผู้วิจัยพัฒนาเครื่องมือวัดขึ้นเอง มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงสุด ($M = .540, SD = .111$) รองลงมาคือ ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยต่างประเทศ ($M = .374, SD = .162$) ผลการเปรียบเทียบค่าเฉลี่ยดัชนีมาตรฐานความยากพบว่า งานวิจัยที่มีผู้วิจัยพัฒนาเครื่องมือวัดขึ้นเอง มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงกว่างานวิจัยที่ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยต่างประเทศ อย่างมีนัยสำคัญทางสถิติ

ตัวแปรประเภทของเครื่องมือวัด พบว่า งานวิจัยที่มีการสร้างเครื่องมือแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงสุด ($M = .693, SD = .080$) รองลงมาคือ อื่น ๆ ($M = .609, SD = .030$), แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า ($M =$

.531, $SD = .092$), แบบสอบอัตนัยประยุกต์ (MEQ) ($M = .530$, $SD = .133$), แบบสอบหลายตัวเลือก ($M = .515$, $SD = .082$), แบบสอบอัตนัย ($M = .512$, $SD = .106$) และแบบสอบภาคปฏิบัติ ($M = .464$, $SD = .157$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยที่มีการสร้างเครื่องมือแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงกว่างานวิจัยที่มีการสร้างเครื่องมือแบบสอบหลายตัวเลือก, แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า, แบบสอบอัตนัย, แบบสอบภาคปฏิบัติ และแบบสอบอัตนัยประยุกต์ (MEQ) อย่างมีนัยสำคัญทางสถิติ

ตัวแปรจำนวนตัวเลือก/เกณฑ์ พบว่า งานวิจัยที่มีการสร้างเครื่องมือที่มี 3 ตัวเลือก/เกณฑ์ มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงสุด ($M = .656$, $SD = .183$) รองลงมาคือ ใช้มากกว่า 1 ประเภท ($M = .618$, $SD = .018$), 5 ตัวเลือก/เกณฑ์ ($M = .610$, $SD = .110$), 4 ตัวเลือก/เกณฑ์ ($M = .527$, $SD = .086$) และไม่มีตัวเลือก/เกณฑ์ ($M = .513$, $SD = .122$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่มีการสร้างเครื่องมือที่มีตัวเลือกมากกว่า 1 ประเภท มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงกว่างานวิจัยที่ไม่มีตัวเลือก/เกณฑ์ และมี 4 ตัวเลือก/เกณฑ์ อย่างมีนัยสำคัญทางสถิติ

ด้านวิวิธนาการวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยากจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านวิวิธนาการวิจัย ประกอบด้วยคุณลักษณะ 2 ตัวแปร ได้แก่ วิธีการเลือกตัวอย่างและจำนวนผู้สอบในแต่ละฉบับ ผลการวิเคราะห์พบว่า ตัวแปรวิธีการเลือกตัวอย่าง เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานความยากแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = .004$) กล่าวคือ เมื่อพิจารณาข้อมูลพื้นฐาน พบว่า งานวิจัยที่ใช้การสุ่มอย่างง่าย มีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงสุด ($M = .697$, $SD = -$) รองลงมาคือ เลือกอย่างเจาะจง ($M = .582$, $SD = .059$), สุ่มแบบหลายขั้นตอน ($M = .574$, $SD = .136$), ไม่ระบุ ($M = .541$, $SD = .120$), ใช้ประชากร/กลุ่มเป้าหมาย ($M = .540$, $SD = .004$), สุ่มแบบแบ่งชั้น ($M = .538$, $SD = .077$) และสุ่มแบบสองขั้นตอน ($M = .425$, $SD = .068$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่ใช้ประชากร/กลุ่มเป้าหมาย, สุ่มแบบแบ่งชั้น, สุ่มแบบหลายขั้นตอน, เลือกอย่างเจาะจง และไม่ระบุวิธีการสุ่ม ค่าเฉลี่ยดัชนีมาตรฐานความยากสูงกว่างานวิจัยที่ใช้สุ่มแบบสองขั้นตอน อย่างมีนัยสำคัญทางสถิติ

ด้านคุณภาพงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยากจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านคุณภาพงานวิจัย มีตัวแปรคุณลักษณะ 1 ตัวแปร คือ ระดับคุณภาพงานวิจัย ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัยด้านคุณภาพงานวิจัยไม่มีผลต่อความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยาก กล่าวคือ งานวิจัยที่มีระดับคุณภาพต่างกันมีค่าเฉลี่ยดัชนีมาตรฐานความยากไม่แตกต่างกัน รายละเอียดดังตารางที่ 4.9

ตารางที่ 4.9 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานความยากจำแนกตามตัวแปรคุณลักษณะงานวิจัย

คุณลักษณะงานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน		test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p
ด้านการพิมพ์และผู้วิจัย								
ปีที่พิมพ์งานวิจัย	(1) พ.ศ. 2546-2550	16	.529	.115	.670	.514	3.568	.032
	(2) พ.ศ. 2551-2555	32	.499	.097				
	(3) พ.ศ. 2556-2561	46	.566	.117				
ประเภทสถาบันที่ผลิตงานวิจัย	(1) มหาวิทยาลัยของรัฐ	18	.540	.085	1.775	.158	2.608	.056
	(2) มหาวิทยาลัยในกำกับของรัฐ	47	.564	.128				
	(3) มหาวิทยาลัยราชภัฏ	24	.494	.096				
	(4) สทศ.	5	.479	.066				
สถาบันที่ผลิตงานวิจัย	(1) จุฬาลงกรณ์มหาวิทยาลัย	30	.594	.137	1.946	.045	1.341	.218
	(2) ม.เกษตรศาสตร์	2	.540	.004				
	(3) ม.เชียงใหม่	4	.504	.133				
	(4) ม.ขอนแก่น	3	.512	.139				
	(5) ม.ทักษิณ	9	.485	.068				
	(6) ม.นเรศวร	3	.531	.137				
	(7) ม.บูรพา	2	.452	.116				
	(8) ม.มหาสารคาม	9	.532	.093				
	(9) ม.สุโขทัยธรรมมาธิราช	7	.552	.046				
	(10) มรภ.กาญจนบุรี	1*	.558	-				
	(11) มรภ.นครราชสีมา	1*	.509	-				
	(12) มรภ.มหาสารคาม	15	.498	.070				
	(13) มรภ.อุบลราชธานี	3	.515	.243				
	(14) สทศ.	5	.479	.066				
สาขาที่ผลิตงานวิจัย	(1) กลุ่มสาขาการวัดและประเมินผล	63	.555	.118	.817	.488	1.900	.135
	(2) กลุ่มสาขาการวิจัยและสถิติ	24	.508	.100				
	(3) เทคโนโลยีการวัดทางการศึกษา	2	.452	.116				
	(4) รายงานวิจัย	5	.479	.066				
ประเภทงานวิจัย	(1) วิทยานิพนธ์ ปริญญาโท	75	.526	.115	1.366	.260	4.774	.011
	(2) วิทยานิพนธ์ ปริญญาเอก	14	.616	.084				
	(3) รายงานวิจัย	5	.479	.066				
เพศของผู้วิจัย	(1) ชาย	6	.546	.084	.815	.446	.122	.885
	(2) หญิง	86	.537	.117				
	(3) ชายและหญิง	2	.500	.057				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
ด้านเนื้อหาสาระงานวิจัย								
องค์ประกอบ ที่มุ่งวัด	(1) ความรู้เรื่องพลเมือง	1*	.588	-	.798	.591	1.895	.081
	(2) การคิดอย่างมี วิจารณญาณ	12	.513	.086				
	(3) การแก้ปัญหา	29	.535	.125				
	(4) การสร้างสรรค์และ นวัตกรรม	13	.464	.109				
	(5) การสื่อสาร	25	.538	.108				
	(6) ความรับผิดชอบ	3	.571	.129				
	(7) การรู้เท่าทันสื่อและ สารสนเทศ	3	.623	.100				
	(8) การรู้เท่าทันเทคโนโลยี	4	.626	.062				
	(9) การคิดอย่างมี วิจารณญาณและการ แก้ปัญหา	2	.656	.008				
	(10) การสื่อสารและการ ร่วมมือ	1*	.639	-				
	(11) ภาวะผู้นำและความ รับผิดชอบ	1*	.669	-				
เนื้อหาที่ เกี่ยวข้อง	(1) คณิตศาสตร์	29	.557	.088	1.636	.173	4.422	.003
	(2) วิทยาศาสตร์	7	.453	.118	ไม่พบคู่ที่มีค่าดัชนีมาตรฐานแตกต่างกัน			
	(3) สังคมศึกษา ศาสนา และวัฒนธรรม*	1*	.352	-				
	(4) ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์)	2	.376	.055				
	(5) การงานอาชีพและ เทคโนโลยี*	1*	.601	-				
	(6) ภาษาต่างประเทศ	7	.446	.044				
	(7) ไม่สังกัดกลุ่มสาระ (content free)	47	.560	.118				
ภูมิภาค	(1) กทม.	9	.620	.161	1.274	.278	2.179	.053
	(2) ภาคเหนือ	12	.534	.100				
	(3) ภาคตะวันออกเฉียงเหนือ	32	.512	.099				
	(4) ภาคกลาง	14	.524	.125				
	(5) ภาคใต้	11	.493	.063				
	(6) มากกว่า 1 ภูมิภาค	14	.596	.115				
	(7) ไม่ระบุ	2	.500	.057				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
แหล่งที่มาของ ตัวอย่าง	(1) สำนักงานคณะกรรมการ การศึกษาขั้นพื้นฐาน (สพฐ.)	81	.543	.114	1.546	.208	1.531	.212
	(2) สำนักงานคณะกรรมการ ส่งเสริมการศึกษาเอกชน*	1*	.549	-				
	(3) อื่น ๆ	2	.540	.004				
	(4) มากกว่า 1 สังกัด	2	.591	.043				
	(5) ไม่ระบุ	8	.458	.120				
ระดับชั้นของ ตัวอย่าง	(1) ชั้นมัธยมศึกษาปีที่ 1	2	.656	.192	1.065	.395	1.854	.078
	(2) ชั้นมัธยมศึกษาปีที่ 2	17	.485	.122				
	(3) ชั้นมัธยมศึกษาปีที่ 3	26	.518	.095				
	(4) ชั้นมัธยมศึกษาปีที่ 4	10	.536	.089				
	(5) ชั้นมัธยมศึกษาปีที่ 5	9	.583	.150				
	(6) ชั้นมัธยมศึกษาปีที่ 6	2	.442	.099				
	(7) ชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน)	22	.586	.093				
	(8) ชั้นมัธยมศึกษาตอนต้น และตอนปลาย	4	.510	.156				
	(9) ชั้นประถมศึกษาและ มัธยมศึกษา	2	.500	.057				
การศึกษา แนวคิดการ สร้างเครื่องมือ วัด	(1) มี เฉพาะการวิเคราะห์ คุณภาพเครื่องมือ	7	.473	.078	.616	.543	1.225	.299
	(2) มี เฉพาะวิธีการสร้าง เครื่องมือ	12	.535	.121				
	(3) มี ทั้งวิธีการสร้างและ การวิเคราะห์คุณภาพ	75	.543	.114				
การสร้าง เครื่องมือวัด	(1) ผู้วิจัยพัฒนาเครื่องมือวัด ขึ้นเอง	92	.540	.111	.331	.566	4.388	.039
	(2) ผู้วิจัยปรับปรุงเครื่องมือ วัดจากงานวิจัยต่างประเทศ	2	.374	.162				
การกำหนด นियามเชิง ปฏิบัติการ	(1) ไม่มี	5	.482	.082	.518	.598	1.138	.325
	(2) มี แต่ไม่นำเสนอ	7	.498	.094				
	(3) มี และนำเสนอ	82	.543	.116				
การสร้าง ตารางกำหนด โครงสร้างของ เครื่องมือวัด	(1) ไม่มี	11	.563	.075	1.711	.186	.471	.626
	(2) มี แต่ไม่นำเสนอ	4	.563	.160				
	(3) มี และนำเสนอ	79	.532	.116				
ประเภทของ เครื่องมือวัด	(1) แบบสอบหลายตัวเลือก	22	.515	.082	1.091	.374	4.301	.001
	(2) แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า	15	.531	.092				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
	(3) แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า	9	.693	.080				
	(4) แบบสอบอัตนัย	26	.512	.106				
	(5) แบบสอบภาคปฏิบัติ	3	.464	.157				
	(6) แบบสอบอัตนัยประยุกต์ (MEQ)	17	.530	.133				
	(7) อื่น ๆ	2	.609	.030				
จำนวน ตัวเลือก/ เกณฑ์	(1) ไม่มีตัวเลือก/เกณฑ์	38	.513	.122	3.084	.020	3.045	.021
	(2) 3 ตัวเลือก/เกณฑ์	4	.656	.183				
	(3) 4 ตัวเลือก/เกณฑ์	41	.527	.086				
	(4) 5 ตัวเลือก/เกณฑ์	9	.610	.110				
	(5) ใช้มากกว่า 1 ประเภท	2	.618	.018				
ลักษณะวิธี ตอบของ เครื่องมือวัด	(1) แบบเขียนตอบ	85	.537	.113	.592	.555	1.760	.161
	(2) คอมพิวเตอร์	6	.562	.096				
	(3) ปากเปล่า	2	.376	.055				
	(4) อื่น ๆ	1*	.640	-				
การตรวจให้ คะแนน	(1) ให้คะแนนแบบสองค่า	47	.525	.086	4.063	.020	.726	.487
	(2) ให้คะแนนมากกว่าสอง ค่า	41	.553	.138				
	(3) ให้คะแนนแบบสองค่า และมากกว่าสองค่า	6	.520	.125				
จำนวนครั้งที่ ใช้ในการ ทดลอง	(1) 1 ครั้ง	4	.479	.131	.468	.706	2.071	.110
	(2) 2 ครั้ง	49	.563	.110				
	(3) 3 ครั้ง	37	.509	.114				
	(4) 4 ครั้ง	4	.528	.074				
จำนวนข้อ คำถามในแต่ละ ฉบับ	(1) น้อยกว่าหรือเท่ากับ 30 ข้อ	83	.541	.117	1.356	.247	1.220	.272
	(2) มากกว่า 30 ข้อ ขึ้นไป	11	.501	.080				
จำนวน ผู้เชี่ยวชาญใน การตรวจสอบ ความตรงตาม เนื้อหา	(1) น้อยกว่าหรือเท่ากับ 5 คน	56	.529	.095	7.609	.007	.661	.418
	(2) มากกว่า 5 คน ขึ้นไป	38	.548	.137				
วิธีที่ใช้ในการ สร้างเกณฑ์ ปกติ	(1) วิธีเปอร์เซ็นต์ไทล์	1*	.436	-	1.299	.280	.927	.431
	(2) วิธีคะแนนมาตรฐานที่ ปกติ	39	.530	.124				
	(3) วิธีเปอร์เซ็นต์ไทล์และ คะแนนมาตรฐานที่รวมกัน	3	.636	.015				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p	
คู่มือการใช้ เครื่องมือวัด	(4) วิธีเปอร์เซ็นต์ไทล์และสเตรนจ์รวมกัน	3	.502	.119	.578	.449	3.567	.062	
	(5) ไม่มีการสร้างเกณฑ์ปกติ	48	.540	.107					
ด้านวิธีวิทยาการวิจัย	(1) มี	18	.582	.125	3.303	.009	3.757	.004	
	(2) ไม่มี	76	.526	.109					
	วิธีการเลือกตัวอย่าง	(1) ใช้ประชากร/กลุ่มเป้าหมาย	2	.540					.004
	(2) สุ่มอย่างง่าย	1*	.697	-					(1), (3), (5), (6) และ (7) > (4)
	(3) สุ่มแบบแบ่งชั้น	18	.538	.077					
	(4) สุ่มแบบสองขั้นตอน	13	.425	.068					
	(5) สุ่มแบบหลายขั้นตอน	19	.574	.136					
(6) เลือกอย่างเจาะจง	9	.582	.059						
(7) ไม่ระบุ	32	.541	.120						
จำนวนผู้สอบ ในแต่ละฉบับ	(1) น้อยกว่าหรือเท่ากับ 200 คน	44	.519	.111	.000	.994	2.006	.160	
	(2) มากกว่า 200 คน ขึ้นไป	50	.552	.114					
ด้านคุณภาพงานวิจัย	(1) ดี	72	.524	.106	1.772	.186	3.774	.055	
	(2) ดีมาก	22	.577	.130					

* ไม่ได้นำเข้าร่วมวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA)

3.2.4 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกจำแนกตามตัวแปรคุณลักษณะงานวิจัย

ด้านการพิมพ์และผู้วิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านการพิมพ์และผู้วิจัย ประกอบด้วยคุณลักษณะ 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, ประเภทงานวิจัย และเพศของผู้วิจัย ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัย ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย และเพศของผู้วิจัย เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p=.000-.041$) มีรายละเอียดดังนี้

ตัวแปรปีที่พิมพ์งานวิจัย พบว่า งานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2546-2550 มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .495, SD = .151$) รองลงมาคือ งานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2556-2561 ($M = .467, SD = .130$) และ พ.ศ. 2551-2555 ($M = .418, SD = .133$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2546-2550 มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยที่มีการตีพิมพ์ในปี พ.ศ. 2551-2555 อย่างมีนัยสำคัญทางสถิติ

ตัวแปรประเภทสถาบันที่ผลิตงานวิจัย พบว่า งานวิจัยจากมหาวิทยาลัยราชภัฏ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .521, SD = .128$) รองลงมาคือ สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ($M = .498, SD = .108$), มหาวิทยาลัยในกำกับของรัฐ ($M = .429, SD = .118$) และมหาวิทยาลัยของรัฐ ($M = .403, SD = .194$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยจากมหาวิทยาลัยราชภัฏ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากมหาวิทยาลัยในกำกับของรัฐอย่างมีนัยสำคัญทางสถิติ

ตัวแปรสถาบันที่ผลิตงานวิจัย พบว่า งานวิจัยจากมหาวิทยาลัยราชภัฏเลย มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .665, SD = .071$) รองลงมาคือ มหาวิทยาลัยสงขลานครินทร์ ($M = .601, SD = .147$), มหาวิทยาลัยราชภัฏอุบลราชธานี ($M = .575, SD = .143$), มหาวิทยาลัยราชภัฏมหาสารคาม ($M = .503, SD = .117$), สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) ($M = .498, SD = .108$), มหาวิทยาลัยทักษิณ ($M = .494, SD = .101$), มหาวิทยาลัยเชียงใหม่ ($M = .477, SD = .097$), มหาวิทยาลัยราชภัฏกาญจนบุรี ($M = .477, SD = .100$), มหาวิทยาลัยเกษตรศาสตร์ ($M = .431, SD = .044$), จุฬาลงกรณ์มหาวิทยาลัย ($M = .431, SD = .132$), มหาวิทยาลัยราชภัฏนครราชสีมา ($M = .431, SD = .132$), มหาวิทยาลัยบูรพา ($M = .396, SD = .169$), มหาวิทยาลัยศรีนครินทรวิโรฒ ($M = .388, SD = .046$), มหาวิทยาลัยสุโขทัยธรรมาธิราช ($M = .378, SD = .008$), มหาวิทยาลัยนเรศวร ($M = .340, SD = .146$), มหาวิทยาลัยขอนแก่น ($M = .289, SD = .026$) และมหาวิทยาลัยมหาสารคาม ($M = .276, SD = .082$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยจากจุฬาลงกรณ์มหาวิทยาลัย, มหาวิทยาลัยเกษตรศาสตร์, มหาวิทยาลัยทักษิณ, มหาวิทยาลัยสุโขทัยธรรมาธิราช, มหาวิทยาลัยราชภัฏมหาสารคาม, มหาวิทยาลัยราชภัฏอุบลราชธานี และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากมหาวิทยาลัยขอนแก่น และมหาวิทยาลัยมหาสารคาม อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ ยังพบว่า งานวิจัยจาก มหาวิทยาลัยราชภัฏมหาสารคาม, มหาวิทยาลัยราชภัฏอุบลราชธานี และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากมหาวิทยาลัยราชภัฏกาญจนบุรี อย่างมีนัยสำคัญทางสถิติ รวมถึง งานวิจัยจากมหาวิทยาลัยราชภัฏเลยมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากมหาวิทยาลัยมหาสารคาม และงานวิจัยจากมหาวิทยาลัยราชภัฏอุบลราชธานีมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากมหาวิทยาลัยศรีนครินทรวิโรฒ อย่างมีนัยสำคัญทางสถิติ

ตัวแปรสาขาที่ผลิตงานวิจัย พบว่า งานวิจัยจากกลุ่มสาขาการวิจัยและสถิติ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .506, SD = .126$) รองลงมาคือ รายงานวิจัย ($M = .498, SD = .108$), กลุ่มสาขาการวัดและประเมินผล ($M = .426, SD = .140$) และเทคโนโลยีการวัดทางการศึกษา ($M = .298, SD = .006$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยจากกลุ่มสาขาการวัดและประเมินผล, กลุ่มสาขาการวิจัยและสถิติ และรายงานวิจัย มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากเทคโนโลยีการวัดทางการศึกษา อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยจากกลุ่มสาขาการวิจัยและสถิติมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากกลุ่มสาขาการวัดและประเมินผล อย่างมีนัยสำคัญทางสถิติ

ตัวแปรเพศของผู้วิจัย พบว่า งานวิจัยจากนักวิจัยชาย มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .548, SD = .125$) รองลงมาคือ งานวิจัยจากนักวิจัยหญิง ($M = .443, SD = .134$) และงานวิจัยจากนักวิจัยชายและหญิง ($M = .415, SD = .049$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย

Bonferroni พบว่า งานวิจัยจากนักวิจัยชาย และงานวิจัยจากนักวิจัยหญิงมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยจากชายและหญิง อย่างมีนัยสำคัญทางสถิติ

ด้านเนื้อหาสาระงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกจำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านเนื้อหาสาระงานวิจัย ประกอบด้วยคุณลักษณะ 18 ตัวแปร ได้แก่ องค์ประกอบที่มุ่งวัด, เนื้อหาที่เกี่ยวข้อง, ภูมิภาค, แหล่งที่มาของตัวอย่าง, ระดับชั้นของตัวอย่าง, การศึกษาแนวความคิด, สร้างเครื่องมือวัด, การกำหนดนิยามเชิงปฏิบัติการ, การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, ลักษณะวิธีตอบของเครื่องมือวัด, การตรวจให้คะแนน, จำนวนครั้งที่ใช้ในการทดลอง, จำนวนข้อคำถามในแต่ละฉบับ, จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา, วิธีที่ใช้ในการสร้างเกณฑ์ปกติ และคู่มือการใช้เครื่องมือวัด ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัย ได้แก่ ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนข้อคำถามในแต่ละฉบับ และวิธีที่ใช้ในการสร้างเกณฑ์ปกติ เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p=.000-.036$) มีรายละเอียดดังนี้

ตัวแปรภูมิภาค พบว่า งานวิจัยที่มีการศึกษาในภาคใต้ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .537, SD = .100$) รองลงมาคือ งานวิจัยที่มีการศึกษาในภาคเหนือ ($M = .518, SD = .140$), มากกว่า 1 ภูมิภาค ($M = .475, SD = .113$), ภาคตะวันออกเฉียงเหนือ ($M = .474, SD = .159$), ไม่ระบุภูมิภาค ($M = .415, SD = .049$), ภาคกลาง ($M = .415, SD = .095$) และกรุงเทพมหานคร ($M = .343, SD = .094$), ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยที่มีการศึกษาในภาคเหนือ, ภาคตะวันออกเฉียงเหนือ, ภาคใต้ และมากกว่า 1 ภูมิภาค มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยที่มีการศึกษาในกรุงเทพมหานคร อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยที่มีการศึกษาในภาคใต้ค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยที่มีการศึกษาในภาคกลาง อย่างมีนัยสำคัญทางสถิติ

ตัวแปรระดับชั้นของตัวอย่าง พบว่า งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้นมัธยมศึกษาปีที่ 1 มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .612, SD = .134$) รองลงมาคือ งานวิจัยที่มีการศึกษาในระดับชั้นมัธยมศึกษาตอนต้นและตอนปลาย ($M = .519, SD = .086$), ชั้นมัธยมศึกษาปีที่ 3 ($M = .504, SD = .132$), ชั้นประถมศึกษาและมัธยมศึกษา ($M = .488, SD = .071$), ชั้นมัธยมศึกษาตอนต้น ($M = .444, SD = .125$), ชั้นมัธยมศึกษาปีที่ 2 ($M = .440, SD = .113$), ชั้นมัธยมศึกษาปีที่ 5 ($M = .416, SD = .122$), ชั้นมัธยมศึกษาตอนปลาย ($M = .395, SD = .096$), ชั้นมัธยมศึกษาปีที่ 6 ($M = .314, SD = .066$) และชั้นมัธยมศึกษาปีที่ 4 ($M = .261, SD = .046$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้นมัธยมศึกษาปีที่ 1 มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่า งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้นมัธยมศึกษาปีที่ 2, 4, 5, 6, ชั้นมัธยมศึกษาตอนต้น และชั้นมัธยมศึกษาตอนปลาย อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้นมัธยมศึกษาปีที่ 2, 3, ชั้นมัธยมศึกษาตอนต้น, ชั้นมัธยมศึกษาตอนต้นและตอนปลาย และชั้นประถมศึกษาและมัธยมศึกษา มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่า งานวิจัยที่มีการศึกษากับตัวอย่างในระดับชั้นมัธยมศึกษาปีที่ 4 อย่างมีนัยสำคัญทางสถิติ

ตัวแปรประเภทของเครื่องมือวัด พบว่า งานวิจัยที่มีการสร้างเครื่องมือแบบสอบอัตโนมัติ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .495, SD = .162$) รองลงมาคือ มาตรฐานประมาณค่า ($M = .493, SD = .096$), อื่น ๆ ($M = .480, SD = .188$), แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า ($M = .474, SD = .114$), แบบสอบหลายตัวเลือก ($M = .457, SD = .168$), แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า ($M = .448, SD = .096$), แบบสอบภาคปฏิบัติ ($M = .391, SD = .018$) และแบบสอบอัตโนมัติประยุกต์ (MEQ) ($M = .362, SD = .135$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Dunnett's T3 พบว่า งานวิจัยมีการสร้างเครื่องมือแบบสอบอัตโนมัติ และมาตรฐานประมาณค่า มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าแบบสอบภาคปฏิบัติ อย่างมีนัยสำคัญทางสถิติ นอกจากนี้ งานวิจัยมีการสร้างเครื่องมือมาตรฐานประมาณค่าค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าแบบสอบอัตโนมัติประยุกต์ (MEQ) อย่างมีนัยสำคัญทางสถิติ

ตัวแปรจำนวนข้อคำถามในแต่ละฉบับ พบว่า งานวิจัยที่มีจำนวนข้อคำถามในแต่ละฉบับน้อยกว่าหรือเท่ากับ 30 ข้อ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .469, SD = .137$) รองลงมาคือ มากกว่า 30 ข้อ ขึ้นไป ($M = .380, SD = .115$) ผลการเปรียบเทียบค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนก พบว่า งานวิจัยที่มีจำนวนข้อคำถามในแต่ละฉบับน้อยกว่าหรือเท่ากับ 30 ข้อ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่า งานวิจัยที่มีจำนวนข้อคำถามในแต่ละฉบับมากกว่า 30 ข้อ ขึ้นไป อย่างมีนัยสำคัญทางสถิติ

ตัวแปรวิธีที่ใช้ในการสร้างเกณฑ์ปกติ พบว่า งานวิจัยที่สร้างเกณฑ์ปกติด้วยวิธีคะแนนมาตรฐานที่ปกติ มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .491, SD = .127$) รองลงมาคือ งานวิจัยที่ไม่มีการสร้างเกณฑ์ปกติ ($M = .434, SD = .148$), วิธีเปอร์เซ็นต์ไทล์ ($M = .396, SD = .061$), วิธีเปอร์เซ็นต์ไทล์และคะแนนมาตรฐานที่รวมกัน ($M = .384, SD = .109$) และวิธีเปอร์เซ็นต์ไทล์และสแตนด์แนลรวมกัน ($M = .324, SD = .045$) ตามลำดับ ผลการทดสอบค่าเฉลี่ยรายคู่ด้วย Bonferroni พบว่า ไม่พบคู่ที่มีค่าดัชนีมาตรฐานแตกต่างกัน

ด้านวิธีวิทยาการวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนก จำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านวิธีวิทยาการวิจัย ประกอบด้วยคุณลักษณะ 2 ตัวแปร ได้แก่ วิธีการเลือกตัวอย่าง และจำนวนผู้สอบในแต่ละฉบับ ผลการวิเคราะห์พบว่า ตัวแปรจำนวนผู้สอบในแต่ละฉบับ เป็นตัวแปรคุณลักษณะหรือตัวแปรปรับที่ทำให้ค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 ($p = .030$) กล่าวคือ เมื่อพิจารณาข้อมูลพื้นฐาน พบว่า งานวิจัยที่มีจำนวนผู้สอบในแต่ละฉบับมากกว่า 200 คน ขึ้นไป มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงที่สุด ($M = .478, SD = .125$) รองลงมาคือ งานวิจัยที่มีจำนวนผู้สอบในแต่ละฉบับน้อยกว่าหรือเท่ากับ 200 คน ($M = .429, SD = .152$) ผลการเปรียบเทียบค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนก พบว่า งานวิจัยที่มีจำนวนผู้สอบในแต่ละฉบับมากกว่า 200 คน ขึ้นไป มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่างานวิจัยที่มีจำนวนผู้สอบในแต่ละฉบับน้อยกว่าหรือเท่ากับ 200 คน อย่างมีนัยสำคัญทางสถิติ

ด้านคุณภาพงานวิจัย

การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนก จำแนกตามตัวแปรคุณลักษณะงานวิจัยด้านคุณภาพงานวิจัย มีตัวแปรคุณลักษณะ 1 ตัวแปร คือ ระดับคุณภาพงานวิจัย ผลการวิเคราะห์พบว่า ตัวแปรคุณลักษณะงานวิจัยด้านคุณภาพงานวิจัยไม่มีผลต่อความแตกต่างของ

ค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนก กล่าวคือ งานวิจัยที่มีระดับคุณภาพต่างกันมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกไม่แตกต่างกัน รายละเอียดดังตารางที่ 4.10

ตารางที่ 4.10 การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกจำแนกตามตัวแปรคุณลักษณะงานวิจัย

คุณลักษณะงานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน		test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p
ด้านการพิมพ์และผู้วิจัย								
ปีที่พิมพ์งานวิจัย	(1) พ.ศ. 2546-2550	34	.495	.151	1.117	.330	3.267	.041
	(2) พ.ศ. 2551-2555	43	.418	.133			(1) > (2)	
	(3) พ.ศ. 2556-2561	84	.467	.130				
ประเภทสถาบันที่ผลิตงานวิจัย	(1) มหาวิทยาลัยของรัฐ	19	.403	.194	7.876	.000	6.488	.000
	(2) มหาวิทยาลัยในกำกับของรัฐ	79	.429	.118			(3) > (2)	
	(3) มหาวิทยาลัยราชภัฏ	46	.521	.128				
	(4) สทศ.	17	.498	.108				
สถาบันที่ผลิตงานวิจัย	(1) จุฬาลงกรณ์มหาวิทยาลัย	44	.431	.132	2.262	.007	5.466	.000
	(2) ม.เกษตรศาสตร์	12	.431	.044			(1), (2), (5), (11), (14), (16) และ (17) > (4) และ (8)	
	(3) ม.เชียงใหม่	4	.477	.097			(14), (16) และ (17) > (12)	
	(4) ม.ขอนแก่น	3	.289	.026			(15) > (8)	
	(5) ม.ทักษิณ	10	.494	.101			(16) > (9)	
	(6) ม.นเรศวร	3	.340	.146				
	(7) ม.บูรพา	3	.396	.169				
	(8) ม.มหาสารคาม	10	.276	.082				
	(9) ม.ศรีนครินทรวิโรฒ	6	.388	.046				
	(10) ม.สงขลานครินทร์	7	.601	.147				
	(11) ม.สุโขทัยธรรมมาธิราช	2	.378	.008				
	(12) มรภ.กาญจนบุรี	3	.477	.100				
	(13) มรภ.นครราชสีมา	44	.431	.132				
	(14) มรภ.มหาสารคาม	21	.503	.117				
	(15) มรภ.เลย	3	.665	.071				
	(16) มรภ.อุบลราชธานี	13	.575	.143				
	(17) สทศ.	17	.498	.108				
สาขาที่ผลิตงานวิจัย	(1) กลุ่มสาขาการวัดและประเมินผล	86	.426	.140	1.777	.154	5.743	.001
	(2) กลุ่มสาขาการวิจัยและสถิติ	56	.506	.126			(1), (2) และ (4) > (3) (2) > (1)	
	(3) เทคโนโลยีการวัดทางการศึกษา	2	.298	.006				
	(4) รายงานวิจัย	17	.498	.108				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
ประเภท งานวิจัย	(1) วิทยานิพนธ์ ปริญญาโท	123	.451	.143	1.892	.154	1.198	.305
	(2) วิทยานิพนธ์ ปริญญาเอก	21	.481	.123				
	(3) รายงานวิจัย	17	.498	.108				
เพศของผู้วิจัย	(1) ชาย	26	.548	.125	.972	.381	6.831	.001
	(2) หญิง	133	.443	.134				
	(3) ชายและหญิง	2	.415	.049				
ด้านเนื้อหาสาระงานวิจัย								
องค์ประกอบ ที่มุ่งวัด	(1) จิตสำนึกต่อโลก	3	.335	.052	2.009	.024	1.406	.163
	(2) ความรู้เรื่องพลเมือง	2	.522	.192				
	(3) การคิดอย่างมี วิจาร์ณญาณ	13	.403	.165				
	(4) การแก้ปัญหา	35	.415	.156				
	(5) การสร้างสรรค์และ นวัตกรรม	33	.509	.153				
	(6) การสื่อสาร	31	.448	.128				
	(7) ทักษะด้านการทำงานเป็น ทีม ความร่วมมือ	5	.499	.112				
	(8) ภาวะผู้นำ	3	.549	.038				
	(9) ความรับผิดชอบ	12	.465	.092				
	(10) การรู้เท่าทันสื่อและ สารสนเทศ	6	.534	.072				
	(11) การรู้เท่าทันเทคโนโลยี	8	.452	.088				
	(12) การคิดอย่างมี วิจาร์ณญาณและการ แก้ปัญหา	4	.527	.076				
	(13) การแก้ปัญหาย่าง สร้างสรรค์	1*	.497	-				
	(14) การสื่อสารและการ ร่วมมือ	2	.464	.072				
	(15) ภาวะผู้นำและความ รับผิดชอบ	3	.486	.006				
เนื้อหาที่ เกี่ยวข้อง	(1) คณิตศาสตร์	36	.493	.178	4.948	.000	.884	.493
	(2) วิทยาศาสตร์	15	.449	.141				
	(3) สังคมศึกษา ศาสนา และวัฒนธรรม*	1*	.311	-				
	(4) ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์)	2	.387	.023				
	(5) การงานอาชีพและ เทคโนโลยี	1*	.375	-				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน			test of homogeneity of variances		ANOVA	
		n	M	SD	Levene statistic	p	F	p
	(6) ภาษาต่างประเทศ	10	.504	.135				
	(7) ไม่สังกัดกลุ่มสาระ (content free)	94	.449	.120				
	(8) มากกว่า 1 สาระ	2	.437	.011				
ภูมิภาค	(1) กทม.	17	.343	.094	4.206	.001	4.320	.000
	(2) ภาคเหนือ	18	.518	.140				
	(3) ภาคตะวันออกเฉียงเหนือ	57	.474	.159				
	(4) ภาคกลาง	25	.415	.095				
	(5) ภาคใต้	12	.537	.100				
	(6) มากกว่า 1 ภูมิภาค	30	.475	.113				
	(7) ไม่ระบุ	2	.415	.049				
แหล่งที่มาของ ตัวอย่าง	(1) สำนักงานคณะกรรมการ การศึกษาขั้นพื้นฐาน (สพฐ.)	131	.454	.141	2.418	.051	.603	.661
	(2) สำนักงานคณะกรรมการ ส่งเสริมการศึกษาเอกชน	1*	.554	-				
	(3) สังกัดองค์กรปกครอง ส่วนท้องถิ่น	1*	.605	-				
	(4) สังกัดคณะกรรมการการ อุดมศึกษา (สกอ.)	3	.426	.033				
	(5) อื่น ๆ	2	.456	.030				
	(6) มากกว่า 1 สังกัด	7	.532	.073				
	(7) ไม่ระบุ	16	.469	.145				
ระดับชั้นของ ตัวอย่าง	(1) ชั้นมัธยมศึกษาปีที่ 1	14	.612	.134	1.700	.094	8.058	.000
	(2) ชั้นมัธยมศึกษาปีที่ 2	18	.440	.113				
	(3) ชั้นมัธยมศึกษาปีที่ 3	35	.504	.132				
	(4) ชั้นมัธยมศึกษาปีที่ 4	10	.261	.046				
	(5) ชั้นมัธยมศึกษาปีที่ 5	9	.416	.122				
	(6) ชั้นมัธยมศึกษาปีที่ 6	3	.314	.066				
	(7) ชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน)	42	.444	.125				
	(8) ชั้นมัธยมศึกษาตอน ปลาย (มากกว่า 1 ชั้นเรียน)	11	.395	.096				
	(9) ชั้นมัธยมศึกษาตอนต้น และตอนปลาย	11	.519	.086				
	(10) ชั้นประถมศึกษาและ มัธยมศึกษา	8	.488	.071				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน		test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p
การศึกษา แนวคิดการ สร้างเครื่องมือ วัด	(1) มี เฉพาะการวิเคราะห์ คุณภาพเครื่องมือ	11	.511	.107	1.260	.287	2.255	.108
	(2) มี เฉพาะวิธีการสร้าง เครื่องมือ	34	.491	.161				
	(3) มี ทั้งวิธีการสร้างและ การวิเคราะห์คุณภาพ	116	.446	.131				
การสร้าง เครื่องมือวัด	(1) ผู้วิจัยพัฒนาเครื่องมือวัด ขึ้นเอง	159	.460	.138	.221	.639	.200	.655
	(2) ผู้วิจัยปรับปรุงเครื่องมือ วัดจากงานวิจัยต่างประเทศ	2	.417	.122				
การกำหนด นियามเชิง ปฏิบัติการ	(1) ไม่มี	8	.484	.198	3.108	.047	.457	.634
	(2) มี แต่ไม่นำเสนอ	13	.488	.127				
	(3) มี และนำเสนอ	140	.456	.135				
การสร้าง ตารางกำหนด โครงสร้างของ เครื่องมือวัด	(1) ไม่มี	29	.493	.164	1.896	.153	2.520	.084
	(2) มี แต่ไม่นำเสนอ	6	.545	.109				
	(3) มี และนำเสนอ	126	.448	.130				
ประเภทของ เครื่องมือวัด	(1) แบบสอบหลายตัวเลือก	22	.457	.168	4.540	.000	2.209	.036
	(2) แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า	22	.474	.114			(4), (6) > (5)	
	(3) แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า	30	.448	.096			(6) > (7)	
	(4) แบบสอบอัตนัย	38	.495	.162				
	(5) แบบสอบภาคปฏิบัติ	3	.391	.018				
	(6) มาตรฐานค่า	23	.493	.096				
	(7) แบบสอบอัตนัยประยุกต์ (MEQ)	19	.362	.135				
	(8) อื่น ๆ	4	.480	.188				
จำนวน ตัวเลือก/ เกณฑ์	(1) ไม่มีตัวเลือก/เกณฑ์	51	.490	.152	3.018	.013	1.765	.123
	(2) 3 ตัวเลือก/เกณฑ์	9	.387	.048				
	(3) 4 ตัวเลือก/เกณฑ์	56	.441	.146				
	(4) 5 ตัวเลือก/เกณฑ์	39	.471	.111				
	(5) มากกว่า 5 ตัวเลือก/ เกณฑ์	2	.525	.094				
	(6) ใช้มากกว่า 1 ประเภท	4	.368	.086				
ลักษณะวิธี ตอบของ เครื่องมือวัด	(1) แบบเขียนตอบ	152	.460	.136	2.411	.093	.513	.600
	(2) คอมพิวเตอร์	6	.498	.201				
	(3) ปากเปล่า	2	.387	.023				
	(4) อื่น ๆ	1*	.400	-				

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน		test of homogeneity of variances		ANOVA			
		n	M	SD	Levene statistic	p	F	p	
การตรวจให้ คะแนน	(1) ให้คะแนนแบบสองค่า	53	.429	.153	3.380	.037	2.678	.072	
	(2) ให้คะแนนมากกว่าสอง ค่า	101	.479	.129					
	(3) ให้คะแนนแบบสองค่า และมากกว่าสองค่า	7	.418	.084					
จำนวนครั้งที่ ใช้ในการ ทดลอง	(1) 1 ครั้ง	4	.328	.120	1.176	.324	1.584	.181	
	(2) 2 ครั้ง	79	.477	.131					
	(3) 3 ครั้ง	63	.443	.145					
	(4) 4 ครั้ง	11	.487	.142					
	(5) 5 ครั้ง	4	.444	.045					
จำนวนข้อ คำถามในแต่ละ ฉบับ	(1) น้อยกว่าหรือเท่ากับ 30 ข้อ	145	.469	.137	1.616	.206	6.234	.014	
	(2) มากกว่า 30 ข้อ ขึ้นไป	16	.380	.115					(1) > (2)
จำนวน ผู้เชี่ยวชาญใน การตรวจสอบ ความตรงตาม เนื้อหา	(1) น้อยกว่าหรือเท่ากับ 5 คน	.459	.141	.459	.554	.458	.003	.955	
	(2) มากกว่า 5 คน ขึ้นไป	.461	.131	.461					
วิธีที่ใช้ในการ สร้างเกณฑ์ ปกติ	(1) วิธีเปอร์เซ็นต์ไทล์	5	.396	.061	2.147	.078	3.050	.019	
	(3) วิธีคะแนนมาตรฐานที่ ปกติ	84	.491	.127					ไม่พบคู่ที่มีค่าดัชนีมาตรฐานแตกต่างกัน
	(4) วิธีเปอร์เซ็นต์ไทล์และ คะแนนมาตรฐานที่รวมกัน	3	.384	.109					
	(5) วิธีเปอร์เซ็นต์ไทล์และสเต โนนัรวมกัน	3	.324	.045					
	(6) ไม่มีการสร้างเกณฑ์ปกติ	66	.434	.148					
	(3) วิธีคะแนนมาตรฐานที่ ปกติ	84	.491	.127					
คู่มือการใช้ เครื่องมือวัด	(1) มี	36	.445	.138	.011	.916	.539	.464	
	(2) ไม่มี	125	.464	.137					
ด้านวิธีวิทยาการวิจัย									
วิธีการเลือก ตัวอย่าง	(1) ใช้ประชากร/กลุ่มเป้าหมาย	2	.456	.030	6.552	.000	1.942	.077	
	(2) สุ่มอย่างง่าย	5	.412	.042					
	(3) สุ่มแบบแบ่งชั้น	22	.431	.193					
	(4) สุ่มแบบสองขั้นตอน	20	.451	.084					
	(5) สุ่มแบบหลายขั้นตอน	47	.515	.141					
	(6) เลือกอย่างเจาะจง	9	.447	.095					
	(7) ไม่ระบุ	56	.434	.127					
จำนวนผู้สอบ ในแต่ละฉบับ	(1) น้อยกว่าหรือเท่ากับ 200 คน	60	.429	.152	3.826	.052	4.776	.030	
	(2) มากกว่า 200 คน ขึ้นไป	101	.478	.125					(2) > (1)

คุณลักษณะ งานวิจัย	ค่าของตัวแปร	ดัชนีมาตรฐาน		test of homogeneity of variances		ANOVA		
		n	M	SD	Levene statistic	p	F	p
ด้านคุณภาพงานวิจัย								
ระดับคุณภาพ งานวิจัย	(1) ดี (2) ดีมาก	119 42	.456 .471	.143 .119	2.279	.133	.375	.541

* ไม่ได้นำเข้าร่วมวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA)

3. ผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา พบว่า

3.1 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความเที่ยง มีจำนวน 11 ตัวแปร ได้แก่ ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, เพศของผู้วิจัย, ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, การตรวจให้คะแนน, จำนวนข้อคำถามในแต่ละฉบับ และวิธีการเลือกตัวอย่าง

3.2 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความตรง มีเพียงตัวแปรเดียวคือ เนื้อหาที่เกี่ยวข้อง

3.3 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความยาก มีจำนวน 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทงานวิจัย, การสร้างเครื่องมือวัด ประเภทของเครื่องมือวัด จำนวนตัวเลือก/เกณฑ์ และวิธีการเลือกตัวอย่าง

3.4 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานอำนาจจำแนก มีจำนวน 10 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, เพศของผู้วิจัย, ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนข้อคำถามในแต่ละฉบับ และจำนวนผู้สอบในแต่ละฉบับ

จากผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ไม่พบตัวแปรคุณลักษณะใดที่มีอิทธิพลต่อดัชนีมาตรฐานทั้งความเที่ยง ความตรง ความยาก และอำนาจจำแนก ทั้งนี้ ตัวแปรคุณลักษณะงานวิจัยที่มีอิทธิพลต่อดัชนีมาตรฐานทั้งความเที่ยง ความยาก และอำนาจจำแนก มี 1 ตัวแปร คือ ประเภทของเครื่องมือวัด ในขณะที่คุณลักษณะงานวิจัยที่มีอิทธิพลต่อดัชนีมาตรฐานทั้งความเที่ยงและอำนาจจำแนก มีจำนวนมากที่สุดจำนวน 7 ตัวแปร ได้แก่ ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, เพศของผู้วิจัย, ภูมิภาค, ระดับชั้นของตัวอย่าง และจำนวนข้อคำถามในแต่ละฉบับ ในขณะที่คุณลักษณะงานวิจัยที่มีอิทธิพลต่อดัชนีมาตรฐานทั้งความเที่ยงและความยาก มี 2 ตัวแปร คือ จำนวนตัวเลือก/เกณฑ์ และวิธีการเลือกตัวอย่าง นอกจากนี้ คุณลักษณะงานวิจัยที่มีอิทธิพลต่อดัชนีมาตรฐานทั้งความยากและอำนาจจำแนก มี 1 ตัวแปร คือ ปีที่พิมพ์งานวิจัย รายละเอียดดังตารางที่ 4.11

ตารางที่ 4.11 สรุปคุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

คุณลักษณะงานวิจัย	สรุปผล	
ด้านการพิมพ์และผู้วิจัย		
1) ปีที่พิมพ์งานวิจัย	ความยาว:	พ.ศ. 2556-2561 > พ.ศ. 2546-2550
	อำนาจจำแนก:	พ.ศ. 2546-2550 > พ.ศ. 2551-2555
2) ประเภทสถาบันที่ผลิตงานวิจัย	ความเที่ยง:	มหาวิทยาลัยราชภัฏ > มหาวิทยาลัยในกำกับของรัฐ
	อำนาจจำแนก:	มหาวิทยาลัยราชภัฏ > มหาวิทยาลัยในกำกับของรัฐ
3) สถาบันที่ผลิตงานวิจัย	ความเที่ยง:	ม.นเรศวร > จุฬาลงกรณ์มหาวิทยาลัย ม.นเรศวร, ม.สงขลานครินทร์ และ ม.สุโขทัยธรรมมาธิราช > ม.มหาสารคาม
	อำนาจจำแนก:	จุฬาลงกรณ์มหาวิทยาลัย, ม.เกษตรศาสตร์, ม.ทักษิณ, ม.สุโขทัยธรรมมาธิราช, มรภ.มหาสารคาม, มรภ.อุบลราชธานี และ สทศ. > ม.ขอนแก่น และ ม.มหาสารคาม มรภ.มหาสารคาม, มรภ.อุบลราชธานี และ สทศ. > มรภ.กาญจนบุรี มรภ.เลย > ม.มหาสารคาม มรภ.อุบลราชธานี > ม.ศรีนครินทรวิโรฒ
4) สาขาที่ผลิตงานวิจัย	ความเที่ยง:	กลุ่มสาขาการวิจัยและสถิติ > กลุ่มสาขาการวัดและประเมินผล
	อำนาจจำแนก:	กลุ่มสาขาการวิจัยและสถิติ > กลุ่มสาขาการวัดและประเมินผล กลุ่มสาขาการวิจัยและสถิติ, กลุ่มสาขาการวัดและประเมินผล และ รายงานวิจัย > เทคโนโลยีการวัดทางการศึกษา
5) ประเภทงานวิจัย	ความยาว:	วิทยานิพนธ์ ปริญญาเอก > วิทยานิพนธ์ ปริญญาโท
6) เพศของผู้วิจัย	ความเที่ยง:	ชาย หรือ หญิง > ชายและหญิง
	อำนาจจำแนก:	ชาย หรือ หญิง > ชายและหญิง
ด้านเนื้อหาสาระงานวิจัย		
7) องค์ประกอบที่มุ่งวัด		-
8) เนื้อหาที่เกี่ยวข้อง	ความตรง:	วิทยาศาสตร์ > คณิตศาสตร์, ภาษาอังกฤษต่างประเทศ, ไม่สังกัดกลุ่มสาระ (content free)
9) ภูมิภาค	ความเที่ยง:	กทม., ภาคเหนือ และภาคใต้ > ไม่ระบุ
	อำนาจจำแนก:	ภาคเหนือ, ภาคตะวันออกเฉียงเหนือ, ภาคใต้ และมากกว่า 1 ภูมิภาค > กทม. ภาคใต้ > ภาคกลาง
10) แหล่งที่มาของตัวอย่าง		-
11) ระดับชั้นของตัวอย่าง	ความเที่ยง:	ชั้นมัธยมศึกษาปีที่ 2, 5 และชั้นมัธยมศึกษาตอนต้นและตอนปลาย > ชั้นมัธยมศึกษาปีที่ 4 ชั้นมัธยมศึกษาปีที่ 1, 2, 3, 5 และชั้นมัธยมศึกษาตอนต้นและตอน ปลาย > ชั้นมัธยมศึกษาปีที่ 6 ชั้นมัธยมศึกษาปีที่ 2 และ 3 > ชั้นมัธยมศึกษาตอนต้น
	อำนาจจำแนก:	ชั้นมัธยมศึกษาปีที่ 1 > ชั้นมัธยมศึกษาปีที่ 2, 4, 5, 6, ชั้น มัธยมศึกษาตอนต้น และชั้นมัธยมศึกษาตอนปลาย ชั้นมัธยมศึกษาปีที่ 2, 3, ชั้นมัธยมศึกษาตอนต้น, ชั้นมัธยมศึกษา ตอนต้นและตอนปลาย และชั้นประถมศึกษาและมัธยมศึกษา > ชั้น มัธยมศึกษาปีที่ 4

คุณลักษณะงานวิจัย		สรุปผล
12) การศึกษาแนวคิดการสร้างเครื่องมือวัด	-	-
13) การสร้างเครื่องมือวัด	ความยาก:	ผู้วิจัยพัฒนาเครื่องมือวัดขึ้นเอง > ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยต่างประเทศ
14) การกำหนดนิยามเชิงปฏิบัติการ	-	-
15) การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด	-	-
16) ประเภทของเครื่องมือวัด	ความเที่ยง:	แบบสังเกต > แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า
		แบบสอบอรรถนัย, มาตรฐานค่า และแบบสังเกต > แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า
	ความยาก:	แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า > แบบสอบหลายตัวเลือก, แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า, แบบสอบอรรถนัย, แบบสอบภาคปฏิบัติ และแบบสอบอรรถนัยประยุกต์ (MEQ)
	อำนาจจำแนก:	แบบสอบอรรถนัย, มาตรฐานค่า > แบบสอบภาคปฏิบัติ มาตรฐานค่า > แบบสอบอรรถนัยประยุกต์ (MEQ)
17) จำนวนตัวเลือก/เกณฑ์	ความเที่ยง:	มากกว่า 5 ตัวเลือก/เกณฑ์ > 3 ตัวเลือก/เกณฑ์ และ 4 ตัวเลือก/เกณฑ์
		ไม่มีตัวเลือก/เกณฑ์ > 4 ตัวเลือก/เกณฑ์
	ความยาก:	เครื่องมือที่มีตัวเลือกมากกว่า 1 ประเภท > ไม่มีตัวเลือก/เกณฑ์ และ 4 ตัวเลือก/เกณฑ์
18) ลักษณะวิธีตอบของเครื่องมือวัด	-	-
19) การตรวจให้คะแนน	ความเที่ยง:	ให้คะแนนมากกว่าสองค่า > ให้คะแนนแบบสองค่า
		ให้คะแนนแบบสองค่าและมากกว่าสองค่า > ให้คะแนนแบบสองค่า และให้คะแนนมากกว่าสองค่า
20) จำนวนครั้งที่ใช้ในการทดลอง	-	-
21) จำนวนข้อคำถามในแต่ละฉบับ	ความเที่ยง:	มากกว่า 30 ข้อ ขึ้นไป > น้อยกว่าหรือเท่ากับ 30 ข้อ
	อำนาจจำแนก:	น้อยกว่าหรือเท่ากับ 30 ข้อ > มากกว่า 30 ข้อ ขึ้นไป
22) จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา	-	-
23) วิธีที่ใช้ในการสร้างเกณฑ์ปกติ	-	-
24) คู่มือการใช้เครื่องมือวัด	-	-
ด้านวิธีวิทยาการวิจัย		
25) วิธีการเลือกตัวอย่าง	ความเที่ยง:	ใช้ประชากร/กลุ่มเป้าหมาย > สุ่มแบบแบ่งชั้น, สุ่มแบบสองขั้นตอน, สุ่มแบบหลายขั้นตอน, เลือกอย่างเจาะจง และไม่ระบุ
		สุ่มอย่างง่าย และสุ่มแบบสองขั้นตอน > ไม่ระบุ
	ความยาก:	ใช้ประชากร/กลุ่มเป้าหมาย, สุ่มแบบแบ่งชั้น, สุ่มแบบหลายขั้นตอน, เลือกอย่างเจาะจง และไม่ระบุ > สุ่มแบบสองขั้นตอน
26) จำนวนผู้สอบในแต่ละฉบับ	อำนาจจำแนก:	มากกว่า 200 คน ขึ้นไป > น้อยกว่าหรือเท่ากับ 200 คน
ด้านคุณภาพงานวิจัย		
27) ระดับคุณภาพงานวิจัย	-	-

ตอนที่ 4 ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

การสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ผู้วิจัยนำงานวิจัยจำนวน 92 เล่ม มาใช้ในการสังเคราะห์เชิงคุณภาพด้วยการวิเคราะห์เชิงเนื้อหา ผู้วิจัยแบ่งการนำเสนอออกเป็น 2 ส่วน คือ 4.1) ผลการสังเคราะห์องค์ความรู้ในภาพรวม และ 4.2) ผลการสังเคราะห์จำแนกตามสมรรถนะ มีรายละเอียดดังนี้

4.1 ผลการสังเคราะห์องค์ความรู้ในภาพรวม

ผลการสังเคราะห์องค์ความรู้ในภาพรวม พบว่า การพัฒนาเครื่องมือ ในงานวิจัยแต่ละเล่มมีความคล้ายคลึงกัน โดยผู้วิจัยแบ่งการพัฒนาเครื่องมือ ออกเป็น 3 ขั้นตอนหลักคือ 1) ขั้นตอนการสร้างข้อคำถาม 2) ขั้นตอนการหาคุณภาพ และ 3) ขั้นตอนการสร้างปกติวิสัยและคู่มือการใช้เครื่องมือ ซึ่งมีรายละเอียดดังนี้

ขั้นตอนการสร้างข้อคำถาม

- 1) กำหนดวัตถุประสงค์ของการสร้างเครื่องมือ
- 2) ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ทั้งเอกสารที่เกี่ยวกับตัวแปรที่มุ่งวัด และเครื่องมือที่ใช้วัด
- 3) กำหนดองค์ประกอบและตัวบ่งชี้ ซึ่งได้จากการสังเคราะห์เอกสาร การสัมภาษณ์ผู้เชี่ยวชาญ หรือการใช้เทคนิคเดลฟาย
- 4) สร้างตารางกำหนดโครงสร้างของเครื่องมือ
- 5) สํารวจ สัมภาษณ์ความคิดเห็นกลุ่มเป้าหมายเพื่อนำสถานการณ์ไปสร้างข้อคำถาม
- 6) สร้างข้อคำถาม และเกณฑ์การให้คะแนน

จากขั้นตอนการพัฒนาเครื่องมือทั้ง 6 ขั้นตอน พบว่า งานวิจัยส่วนใหญ่ดำเนินการครบ 5 ขั้นตอน ยกเว้น ขั้นตอนที่ 5 การสำรวจ สัมภาษณ์ความคิดเห็นกลุ่มเป้าหมายเพื่อนำสถานการณ์ไปสร้างข้อคำถาม พบว่า ขั้นตอนนี้ ดำเนินการในเครื่องมือบางประเภทเท่านั้น เช่น แบบวัดเชิงสถานการณ์ประเภทเลือกตอบ เนื่องจาก การสร้างตัวเลือก หรือสถานการณ์ จะอ้างอิงตามความคิด หรือประสบการณ์ของกลุ่มเป้าหมาย

ขั้นตอนการหาคุณภาพ

- 1) หาความตรงตามเนื้อหาโดยผู้เชี่ยวชาญ ซึ่งสามารถพิจารณาจากการวิเคราะห์ความสอดคล้องระหว่างข้อคำถามและนิยามศัพท์หรือวัตถุประสงค์ในการวัด (Item Objectice Congruence: IOC) หรือ การพิจารณาความเหมาะสมของข้อคำถามด้วยแบบประเมินแบบมาตรฐานค่า
- 2) ทดลองใช้เครื่องมือ (try out) จำนวน 1-5 ครั้ง เพื่อตรวจสอบคุณภาพในประเด็นต่อไปนี้
 - 2.1) ความเที่ยง
 - 2.2) ความตรง
 - 2.3) ความยาก
 - 2.4) อำนาจจำแนก
 - 2.5) การตรวจสอบด้านพฤติกรรม การใช้ภาษาและเวลา
 - 2.6) การวิเคราะห์การทำหน้าที่ต่างกันของข้อคำถาม
 - 2.7) การวิเคราะห์ความไม่แปรเปลี่ยน

ในขั้นตอนของการหาคุณภาพเครื่องมือนั้น สามารถเลือกวิธีวิเคราะห์ได้ตามความเหมาะสม ซึ่งการสังเคราะห์งานวิจัยในครั้งนี้ พบว่า ส่วนใหญ่เลือกใช้การวิเคราะห์ด้วยทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) และบางส่วนใช้การวิเคราะห์ด้วยทฤษฎีการทดสอบแนวใหม่ (modern test theory เช่น IRT และ G theory) รวมถึงการวิเคราะห์คุณภาพเครื่องมือโดยใช้หลายทฤษฎี สำหรับจำนวนครั้งของการทดลองใช้เครื่องมือนั้น งานวิจัยแต่ละเรื่องมีการออกแบบที่แตกต่างกัน งานวิจัยบางเล่ม วิเคราะห์คุณภาพเกี่ยวกับความยาก อำนาจจำแนกและความเที่ยง พร้อมกันในการทดลอง 1 ครั้ง งานวิจัยบางเล่มมีกลุ่มทดลองใช้เครื่องมือหลายกลุ่ม เพื่อแยกวิเคราะห์ที่ละด้าน เช่น ทดลองใช้เครื่องมือครั้งที่ 1 เพื่อตรวจสอบด้านพฤติกรรม การใช้ภาษาและเวลา ครั้งที่ 2 เพื่อตรวจสอบความยาก และอำนาจจำแนก ครั้งที่ 3 เพื่อตรวจสอบความเที่ยง และ ครั้งที่ 4 เพื่อตรวจสอบความตรงตามโครงสร้าง

ขั้นตอนการสร้างปกติวิสัยและคู่มือการใช้เครื่องมือ

ขั้นตอนการสร้างปกติวิสัย (norm) เรียกอีกอย่างหนึ่งว่า เกณฑ์ปกติ หรืองานวิจัยบางประเภทใช้คำว่า คะแนนจุดตัด (cutting score) ซึ่งการกำหนดปกติวิสัย เป็นกระบวนการสุดท้ายของการพัฒนาเครื่องมือวัดสมรรถนะ เพื่อใช้เป็นเกณฑ์ในการแปลผลระดับสมรรถนะของนักเรียน ซึ่งกระบวนการสร้างปกติวิสัยจะใช้คะแนนที่ได้จากกลุ่มตัวอย่างกลุ่มสุดท้ายที่ใช้เครื่องมือฉบับสมบูรณ์ เพื่อแปลงคะแนนดังกล่าวให้อยู่ในรูปของคะแนนมาตรฐาน ซึ่งมีหลายวิธี เช่น วิธีเปอร์เซ็นต์ไทล์ วิธีคะแนนมาตรฐานที่ วิธีคะแนนมาตรฐานที่ปกติ และวิธีสเตโน

4.2 ผลการสังเคราะห์จำแนกตามสมรรถนะ

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ผู้วิจัยจำแนกตามสมรรถนะหลัก 4 สมรรถนะ และสมรรถนะย่อยรวมทั้งสิ้น 11 สมรรถนะ ซึ่งมีรายละเอียดดังนี้

สมรรถนะที่ 1 ทักษะและความรู้ในเนื้อหาขั้นสูง

1.1 จิตสำนึกต่อโลก (global awareness)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดจิตสำนึกต่อโลก พบว่า มีงานวิจัยที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดจิตสำนึกต่อโลกจำนวน 3 เรื่อง โดยงานวิจัย 2 เรื่องศึกษาเกี่ยวกับภาวะโลกร้อน และมีเพียง 1 เรื่องที่พัฒนาเครื่องมือวัดการตระหนักรู้ต่อโลกในมุมมองกว้าง ซึ่งมีตัวบ่งชี้ ดังนี้

- ตัวบ่งชี้ที่ 1 ความตระหนักในมุมมองที่แตกต่าง
- ตัวบ่งชี้ที่ 2 ความตระหนักในสภาพปัจจุบันของโลก
- ตัวบ่งชี้ที่ 3 ความตระหนักในความแตกต่างของวัฒนธรรม
- ตัวบ่งชี้ที่ 4 ความรู้ในเรื่องพลวัตของโลก
- ตัวบ่งชี้ที่ 5 ความตระหนักต่อทางเลือกของมนุษย์

สำหรับความตระหนักในภาวะโลกร้อน จัดเป็นความตระหนักในสภาพปัจจุบันของโลก สำหรับแนวคิดทฤษฎีที่นำมาใช้ในการกำหนดตัวบ่งชี้ เกี่ยวกับความตระหนักต่อโลกยึดตามแนวคิดของ Hanvey (1982) ส่วนความตระหนักต่อภาวะโลกร้อนยึดตามแนวคิดของ Stamm et al. (2000) นอกจากนี้ การพัฒนาเครื่องมือวัดจิตสำนึกต่อโลก พบว่า ผู้พัฒนาเครื่องมือมีเป้าหมายในการสร้างทักษะ

นี้โดยตรง ไม่ได้เป็นองค์ประกอบย่อยของทักษะอื่นๆ เครื่องมือที่พัฒนามีรูปแบบเป็นมาตรฐานประมาณค่า 5 ระดับ และแบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า ซึ่งใช้วัดความรู้ในเรื่องพลวัตของโลก รายละเอียดการสังเคราะห์ ดังตาราง ง.1 ภาคผนวก ง

1.2 ความรู้เรื่องพลเมือง (civic literacy)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดความรู้เรื่องพลเมือง พบว่า มีงานวิจัยเพียง 1 เรื่องที่ศึกษาเกี่ยวกับแบบวัดความเป็นพลเมืองดี ซึ่งได้สังเคราะห์คุณลักษณะของความเป็นพลเมืองดีจากคุณลักษณะอ้างอิงประสงค์ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 และคุณลักษณะพลเมืองของประเทศชาติและสังคมโลก ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 กลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม ซึ่งมีตัวบ่งชี้ ดังนี้

- ตัวบ่งชี้ที่ 1 รักชาติศาสน์กษัตริย์
- ตัวบ่งชี้ที่ 2 ซื่อสัตย์สุจริต
- ตัวบ่งชี้ที่ 3 มีวินัย เคารพกฎหมายและกติกาสังคม
- ตัวบ่งชี้ที่ 4 ใฝ่เรียนรู้
- ตัวบ่งชี้ที่ 5 อยู่อย่างพอเพียง
- ตัวบ่งชี้ที่ 6 มุ่งมั่นรับผิดชอบในการทำงาน
- ตัวบ่งชี้ที่ 7 รักความเป็นไทย
- ตัวบ่งชี้ที่ 8 มีจิตสาธารณะ
- ตัวบ่งชี้ที่ 9 มีเหตุผลรับฟังความคิดเห็นของผู้อื่น
- ตัวบ่งชี้ที่ 10 เข้าร่วมกิจกรรมทางการเมืองการปกครอง

โดยเครื่องมือที่พัฒนาขึ้นมีลักษณะเป็นแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า มีลักษณะเป็นเลือกตอบ 4 ตัวเลือก กำหนดให้คะแนนตั้งแต่ 0-3 รายละเอียดการสังเคราะห์ดังตาราง ง.2 ภาคผนวก ง

สมรรถนะที่ 2 ทักษะการคิดขั้นสูง

2.1 การคิดอย่างมีวิจารณญาณ (critical thinking)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการคิดอย่างมีวิจารณญาณ พบว่า มีงานวิจัยที่พัฒนาเครื่องมือวัดการคิดอย่างมีวิจารณญาณโดยตรงจำนวน 9 เรื่อง และการคิดอย่างมีวิจารณญาณ เป็นองค์ประกอบของเครื่องมือประเภทอื่น ซึ่งรายงานอยู่ในงานวิจัยจำนวน 9 เรื่อง ซึ่งส่วนใหญ่เป็นองค์ประกอบของแบบวัดความสามารถด้านการคิดหรือแบบวัดทักษะการคิด แนวคิดที่นำมาใช้ในการพัฒนาเครื่องมือ ส่วนใหญ่มาจากการสังเคราะห์จากหลายแหล่งข้อมูล โดยมีแนวคิดที่สำคัญ คือ แนวคิดการคิดอย่างมีวิจารณญาณของ Ennis (1985) จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้ โดยพิจารณาจากตัวบ่งชี้ที่มีความถี่สูงสุด 5 อันดับแรก ดังนี้

- ตัวบ่งชี้ที่ 1 การสรุปและตัดสินใจเพื่อนำไปสรุปอ้างอิง
- ตัวบ่งชี้ที่ 2 การระบุข้อตกลงเบื้องต้น
- ตัวบ่งชี้ที่ 3 การระบุค่านิยม ความเชื่อ การพิจารณาความน่าเชื่อถือของแหล่งข้อมูลและการสังเกต

- ตัวบ่งชี้ที่ 4 การนิยามปัญหา/ระบุประเด็นปัญหา

ตัวบ่งชี้ที่ 5 ความสามารถด้านการนิรนัย
เครื่องมือที่พัฒนาขึ้น ส่วนใหญ่คือแบบสอบหลายตัวเลือก ที่ให้คะแนนแบบ 2 ค่า คือ ผิด
ได้ 0 คะแนน ถูกได้ 1 คะแนน รายละเอียดการสังเคราะห์ดังตาราง ง.3 ภาคผนวก ง

2.2 การแก้ปัญหา (solving problems)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการแก้ปัญหา พบว่า มี
การพัฒนาเครื่องมือวัดการแก้ปัญหาจำนวนมาก โดยส่วนใหญ่ยึดตามแนวคิดของโพลยา ซึ่งทักษะการ
แก้ปัญหานี้ เป็นองค์ประกอบย่อยของแบบวัดอีกหลายประเภท เช่น แบบวัดทักษะกระบวนการทาง
คณิตศาสตร์ และแบบวัดทักษะชีวิต

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้ โดยพิจารณาจากตัวบ่งชี้ที่มีความถี่
สูงสุด 5 อันดับแรก และเรียงลำดับตามขั้นตอนการแก้ปัญหา ดังนี้

ตัวบ่งชี้ที่ 1 การทำความเข้าใจสถานการณ์ปัญหา รับรู้ปัญหา

ตัวบ่งชี้ที่ 2 การวางแผนอย่างเป็นระบบ กำหนดเป้าหมายหรือแนวทางของผลลัพธ์ที่ต้องการ

ตัวบ่งชี้ที่ 3 แสวงหาแนวทาง/ทางเลือกในการแก้ปัญหอย่างเป็นระบบ

ตัวบ่งชี้ที่ 4 วิเคราะห์ทางเลือกอย่างรอบคอบ และตรวจสอบความเป็นไปได้

ตัวบ่งชี้ที่ 5 การดำเนินการ ลงมือปฏิบัติอย่างถูกต้องและเหมาะสม

เครื่องมือวัดการแก้ปัญหาที่พัฒนาขึ้น ส่วนใหญ่มีลักษณะเป็นแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า
2 ค่า มีตัวเลือกอยู่ระหว่าง 3 - 5 ตัวเลือก รองลงมาคือ แบบสอบหลายตัวเลือก ที่ให้คะแนนแบบ 2 ค่า คือ
ผิดได้ 0 คะแนน ถูกได้ 1 คะแนน มีตัวเลือกอยู่ระหว่าง 4 - 5 ตัวเลือก รายละเอียดการสังเคราะห์ดังตาราง
ง.4 ภาคผนวก ง

2.3 การสร้างสรรค์และนวัตกรรม (creativity and innovation)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการสร้างสรรค์และ
นวัตกรรม พบว่า มีงานวิจัยที่พัฒนาแบบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์/วิทยาศาสตร์ จำนวน 6
เรื่อง และความคิดสร้างสรรค์เป็นองค์ประกอบของเครื่องมือวัดทักษะอื่น เช่น แบบวัดทักษะชีวิต และแบบวัด
จิตวิทยาศาสตร์ โดยการพัฒนาเครื่องมือการสร้างสรรค์และนวัตกรรมส่วนใหญ่ ใช้แนวคิดของ Torrance
และ Guilford

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้การสร้างสรรค์และนวัตกรรม โดย
พิจารณาจากตัวบ่งชี้ที่มีความถี่สูงสุด 5 อันดับแรก ได้ดังนี้

ตัวบ่งชี้ที่ 1 ความคิดแปลกใหม่ การสร้างองค์ความรู้ใหม่

ตัวบ่งชี้ที่ 2 ความคิดริเริ่ม

ตัวบ่งชี้ที่ 3 ความคิดคล่องแคล่ว

ตัวบ่งชี้ที่ 4 ความคิดยืดหยุ่น

ตัวบ่งชี้ที่ 5 การคิดได้อย่างหลากหลาย

เครื่องมือวัดการสร้างสรรค์และนวัตกรรม งานวิจัยส่วนใหญ่ใช้รูปแบบการเขียนตอบ ด้วย
แบบสอบอัตนัย โดยให้คะแนนจำแนกออกเป็น 3 ด้าน คือ ความคิดคล่องแคล่ว ความคิดยืดหยุ่น และ
ความคิดริเริ่ม ซึ่งความคล่องแคล่ว ให้คะแนนจากคำตอบที่ไม่ซ้ำ โดยนับจำนวนคำตอบ สำหรับความ

ยืดหยุ่น ให้คะแนนตามจำนวนกลุ่มที่สามารถจัดหมวดหมู่ได้ และด้านความคิดริเริ่ม ให้คะแนนตามสัดส่วนของความถี่ของคำตอบตามวิธีการของ Cropley ดังนี้

- คำตอบซ้ำกัน 12.01 ขึ้นไปให้ 0 คะแนน
 - คำตอบซ้ำกัน 6.01 – 12.00 ขึ้นไปให้ 1 คะแนน
 - คำตอบซ้ำกัน 3.01 – 6.00 ขึ้นไปให้ 2 คะแนน
 - คำตอบซ้ำกัน 1.01 – 3.00 ขึ้นไปให้ 3 คะแนน
 - คำตอบซ้ำกัน 0.01 – 1.00 ขึ้นไปให้ 4 คะแนน
- รายละเอียดการสังเคราะห์ดังตาราง ง.5 ภาคผนวก ง

สมรรถนะที่ 3 ทักษะต่อตนเองและสังคม

3.1 การสื่อสาร (communications literacy/skill)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการสื่อสาร พบว่างานวิจัยที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดการสื่อสารมีจำนวน 22 เรื่อง โดยทักษะการสื่อสารเป็นทักษะย่อยของเครื่องมือหลายประเภท เช่น เครื่องมือวัดทักษะชีวิต และเครื่องมือวัดทักษะกระบวนการทางคณิตศาสตร์ นอกจากนี้ เครื่องมือวัดทักษะการสื่อสารจะอยู่ในลักษณะของการสื่อสารทางภาษาในรายวิชาภาษาไทยและภาษาอังกฤษ ซึ่งจะพิจารณาในด้านของหลักภาษาและไวยากรณ์ ทั้งนี้ผู้วิจัยพิจารณาสื่อสารทั้งด้านการรับสารและการส่งสาร ดังนั้นงานวิจัยที่นำมาสังเคราะห์จึงประกอบด้วย งานวิจัยที่เกี่ยวกับเครื่องมือวัดการฟัง พูด อ่าน เขียน

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้การสื่อสาร ได้ดังนี้

ตัวบ่งชี้ที่ 1 ความสามารถในการประเมินและเลือกวิธีสื่อสารได้เหมาะสมกับวัฒนธรรมและสถานการณ์

ตัวบ่งชี้ที่ 2 การสื่อสารความคิด มุมมอง ความรู้ความเข้าใจโดยใช้ภาษาพูด ภาษาเขียน และกิริยาท่าทางได้อย่างเหมาะสม

ตัวบ่งชี้ที่ 3 สามารถจับประเด็น และสรุปสาระสำคัญจากการรับสารได้

เครื่องมือที่งานวิจัยพัฒนาขึ้น ส่วนใหญ่คือ แบบสอบถามหลายตัวเลือก ให้คะแนนแบบ 2 ค่า คือ ผิดได้ 0 คะแนน ถูกได้ 1 คะแนน รองลงมาคือ แบบสอบถามอัตนัย ให้คะแนนตามเกณฑ์รูปรีติที่ผู้พัฒนาเครื่องมือกำหนด รายละเอียดการสังเคราะห์ดังตาราง ง.6 ภาคผนวก ง

3.2 ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (collaboration, teamwork)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดทักษะด้านการทำงานเป็นทีม ความร่วมมือ พบว่า มีงานวิจัยที่ศึกษาเรื่องดังกล่าวจำนวน 10 เรื่อง โดยที่ทักษะด้านการทำงานเป็นทีม ความร่วมมือ เป็นเพียงองค์ประกอบหนึ่งของแบบวัดทักษะประเภทอื่น ๆ เช่น มาตรฐานจิตวิทยาศาสตร์ และแบบวัดคุณลักษณะอันพึงประสงค์

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้การทำงานเป็นทีม ความร่วมมือ ได้ดังนี้

ตัวบ่งชี้ที่ 1 ความเต็มใจที่จะทำงานร่วมกับผู้อื่น

ตัวบ่งชี้ที่ 2 รู้จักบทบาทของตนที่ได้รับมอบหมาย

ตัวบ่งชี้ที่ 3 ให้ความร่วมมือกับผู้อื่น และมีทักษะการทำงานร่วมกับผู้อื่น

ตัวบ่งชี้ที่ 4 ยอมรับความคิดเห็นของผู้อื่น

ตัวบ่งชี้ที่ 5 เห็นคุณค่าของการทำงานร่วมกับผู้อื่น

เครื่องมือที่งานวิจัยพัฒนาขึ้น ส่วนใหญ่คือ เป็นแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า มีตัวเลือกอยู่ระหว่าง 4 - 5 ตัวเลือก รองลงมาคือ มาตรฐานค่า 5 ระดับ รายละเอียดการสังเคราะห์ดังตาราง ง.7 ภาคผนวก ง

3.3 ภาวะผู้นำ (leadership)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดภาวะผู้นำ พบว่า มีงานวิจัยที่พัฒนาเครื่องมือวัดภาวะผู้นำจำนวน 5 เรื่อง ทั้งนี้ไม่มีงานวิจัยเล่มใดที่ศึกษาภาวะผู้นำโดยตรง แต่ภาวะผู้นำถือเป็นองค์ประกอบของทักษะอื่น ๆ เช่น ความมีวินัย และคุณลักษณะที่พึงประสงค์ เป็นต้น

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้ภาวะผู้นำ ได้ดังนี้

ตัวบ่งชี้ที่ 1 ความสามารถในการทำให้ผู้อื่นคล้อยตามหรือปฏิบัติตามแนวคิดหรือคำสั่งของตน

ตัวบ่งชี้ที่ 2 มีความเห็นอกเห็นใจผู้อื่น ไม่เห็นแก่ตัวมีน้ำใจต่อผู้อื่น คิดถึงประโยชน์ส่วนรวม

ตัวบ่งชี้ที่ 3 มีความเสียสละ ไม่คิดร้ายผู้อื่น

ตัวบ่งชี้ที่ 4 มีเหตุผลไม่ลำเอียง ปราศจากอคติ

โดยเครื่องมือที่พัฒนาขึ้นส่วนใหญ่คือ มาตรฐานค่า 5 ระดับ และเป็นแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า แบบ 4 ตัวเลือก รายละเอียดการสังเคราะห์ดังตาราง ง.8 ภาคผนวก ง

3.4 ความรับผิดชอบ (responsibility)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดความรับผิดชอบต่อตนเอง พบว่า มีงานวิจัยที่พัฒนาเครื่องมือวัดความรับผิดชอบต่อตนเอง จำนวน 23 เรื่อง ทั้งความรับผิดชอบต่อตนเอง และความรับผิดชอบต่อสังคม ซึ่งความรับผิดชอบต่อตนเอง เป็นองค์ประกอบย่อยของคุณธรรม จริยธรรม ค่านิยมที่พึงประสงค์ จิตวิทยาศาสตร์ ทักษะชีวิต ความมีวินัยในตนเอง และทักษะในศตวรรษที่ 21

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้ความรับผิดชอบต่อตนเอง ได้ดังนี้

ความรับผิดชอบต่อตนเอง

ตัวบ่งชี้ที่ 1 ยอมรับผลของการกระทำและคำพูดของตน

ตัวบ่งชี้ที่ 2 มุ่งมั่น ตั้งใจ ไม่หลีกเลียง

ตัวบ่งชี้ที่ 3 ทำงานที่ได้รับมอบหมายเต็มความสามารถ

ตัวบ่งชี้ที่ 4 ติดตามและปรับปรุงการปฏิบัติหน้าที่ให้ดียิ่งขึ้น

ตัวบ่งชี้ที่ 5 ละเอียด รอบคอบ ความเอาใจใส่

ความรับผิดชอบต่อสังคม

ตัวบ่งชี้ที่ 1 ปฏิบัติตามกฎหมายระเบียบของสังคม

ตัวบ่งชี้ที่ 2 การทำหน้าที่เพื่อประโยชน์ส่วนรวม

ตัวบ่งชี้ที่ 3 ความรู้สึกที่ตนเองเป็นส่วนหนึ่งของสังคม

โดยเครื่องมือที่พัฒนาขึ้นส่วนใหญ่คือ แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า และมาตรฐานค่า 3 - 6 ระดับ รายละเอียดการสังเคราะห์ดังตาราง ง.9 ภาคผนวก ง

สมรรถนะที่ 4 ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี

4.1 การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการรู้เท่าทันสื่อและสารสนเทศ พบว่า มีงานวิจัยที่พัฒนาเครื่องมือวัดการรู้เท่าทันสื่อและสารสนเทศจำนวน 5 เรื่อง ซึ่งงานวิจัยที่ศึกษาทักษะนี้ ได้ระบุถึงที่มาของการรู้เท่าทันสื่อและสารสนเทศจากการศึกษาทักษะในศตวรรษที่ 21 และถือเป็นทักษะใหม่ในยุคนี้

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้การรู้เท่าทันสื่อและสารสนเทศ ได้ดังนี้

ตัวบ่งชี้ที่ 1 การเข้าถึง (access)

ตัวบ่งชี้ที่ 2 การประเมิน (evaluate)

ตัวบ่งชี้ที่ 3 การสร้างสรรค์ (create)

ตัวบ่งชี้ที่ 4 การมีส่วนร่วม (act)

ตัวบ่งชี้ที่ 5 การวิเคราะห์ (analyze)

โดยเครื่องมือที่พัฒนาขึ้นส่วนใหญ่คือ มาตรฐานค่า 5 ระดับ และแบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า แบบ 4 ตัวเลือก รายละเอียดการสังเคราะห์ดังตาราง ง.10 ภาคผนวก ง

4.2 การรู้เท่าทันเทคโนโลยี (technology literacy)

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการรู้เท่าทันเทคโนโลยี พบว่า มีงานวิจัยที่พัฒนาเครื่องมือวัดการรู้เท่าทันเทคโนโลยี จำนวน 7 เรื่อง เช่น แบบวัดความไว้วางใจระหว่างบุคคลในเครือข่ายสังคมออนไลน์ เครื่องมือวัดสมรรถนะผู้เรียนด้านเทคโนโลยีสารสนเทศ และแบบวัดจริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร

จากงานวิจัยที่นำมาสังเคราะห์สามารถสรุปตัวบ่งชี้การรู้เท่าทันเทคโนโลยี ได้ดังนี้

ตัวบ่งชี้ที่ 1 ความรู้ด้านเทคโนโลยีสารสนเทศ

ตัวบ่งชี้ที่ 2 ทักษะกระบวนการด้านเทคโนโลยีสารสนเทศ

ตัวบ่งชี้ที่ 3 เจตคติในด้านเทคโนโลยีสารสนเทศ

โดยเครื่องมือที่พัฒนาขึ้นส่วนใหญ่คือ มาตรฐานค่า 4 - 6 ระดับ และแบบสอบภาคปฏิบัติ ซึ่งเป็นการปฏิบัติการใช้คอมพิวเตอร์ เป็นต้น รายละเอียดการสังเคราะห์ดังตาราง ง.11 ภาคผนวก ง

สรุปเครื่องมือที่นิยมใช้ในแต่ละสมรรถนะ แบ่งออกเป็น 4 กลุ่ม ดังนี้

กลุ่มที่ 1 สมรรถนะที่ใช้เครื่องมือประเภทมาตรฐานค่า ได้แก่ จิตสำนึกต่อโลก ภาวะผู้นำ การรู้เท่าทันสื่อและสารสนเทศ และการรู้เท่าทันเทคโนโลยี

กลุ่มที่ 2 สมรรถนะที่ใช้เครื่องมือประเภทแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า ได้แก่ ความรู้เรื่องพลเมือง การแก้ปัญหา ทักษะด้านการทำงานเป็นทีม ความร่วมมือ และความรับผิดชอบ

กลุ่มที่ 3 สมรรถนะที่ใช้เครื่องมือประเภทแบบสอบหลายตัวเลือก ให้คะแนน 2 ค่า ได้แก่ การคิดอย่างมีวิจารณญาณ และการสื่อสาร

กลุ่มที่ 4 สมรรถนะที่ใช้เครื่องมือประเภทแบบสอบอัตนัย ได้แก่ การสร้างสรรค์และนวัตกรรม

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อ 1) เพื่อประเมินคุณภาพงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา 2) เพื่อศึกษาคุณลักษณะของงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา และ 3) เพื่อสังเคราะห์องค์ความรู้เกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา โดยมีกลุ่มตัวอย่างสำหรับใช้วิเคราะห์ห่อภิมาณ คือ งานวิจัยประเภทวิทยานิพนธ์ระดับบัณฑิตศึกษา ทั้งระดับมหาบัณฑิตและดุษฎีบัณฑิต และงานวิจัยของสถาบันต่าง ๆ ที่ศึกษาเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา โดยพิจารณาตามเกณฑ์ที่ผู้วิจัยได้กำหนดขึ้น คือ เป็นงานวิจัยที่สามารถสืบค้นฉบับสมบูรณ์ (full text) และมีตีพิมพ์เผยแพร่ระหว่างปี พ.ศ. 2546-2562 ได้ เป็นงานวิจัยที่มีวัตถุประสงค์เพื่อพัฒนาเครื่องมือวัดประเภท แบบสอบ แบบวัด และแบบประเมิน โดยไม่นับรวมงานวิจัยประเภทพัฒนารูปแบบการประเมิน รวมถึงเครื่องมือเชิงคุณภาพ และมีการรายงานค่าคุณภาพของเครื่องมือในมิติความเที่ยง ความตรง ความยาก และอำนาจจำแนก ด้วยทฤษฎีการทดสอบแบบดั้งเดิม ค่าใดค่าหนึ่งเป็นอย่างน้อย ทั้งนี้ มีงานวิจัยที่นำมาใช้ในการวิเคราะห์ห่อภิมาณจำนวน 90 เล่ม และงานวิจัยสำหรับการสังเคราะห์เชิงคุณภาพจำนวน 92 เล่ม เครื่องมือที่ใช้ในการวิจัย 2 ฉบับ คือ แบบประเมินคุณภาพงานวิจัย และแบบบันทึกคุณลักษณะงานวิจัย เครื่องมือทั้งสองฉบับมีค่าความสอดคล้องระหว่างผู้ประเมินอยู่ระหว่าง 0.88-0.94 และ 0.96-0.99 ตามลำดับ วิเคราะห์ข้อมูลด้วย 1) วิเคราะห์ข้อมูลพื้นฐาน ด้วยสถิติเชิงบรรยาย ได้แก่ ค่าความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน เพื่อรายงานผลการประเมินคุณภาพงานวิจัย และวิเคราะห์สัมประสิทธิ์การกระจาย ค่าต่ำสุด ค่าสูงสุด ความเบ้ ความโด่งเพื่อให้ทราบลักษณะการแจกแจงการกระจายของตัวแปรคุณลักษณะงานวิจัย และค่าดัชนีมาตรฐาน 2) การวิเคราะห์ความแตกต่างของค่าเฉลี่ยดัชนีมาตรฐานจำแนกตามตัวแปรคุณลักษณะงานวิจัย โดยใช้การวิเคราะห์ความแปรปรวนทางเดียว (one-way ANOVA) และทดสอบเอกพันธ์ของความแปรปรวน (test of homogeneity of variance) ตรวจสอบข้อตกลงเบื้องต้นโดยใช้การทดสอบ Levene (Levene's test) เพื่อทำการทดสอบความแตกต่างของค่าเฉลี่ยรายคู่ ในกรณีที่ความแปรปรวนมีความเป็นเอกพันธ์ผู้วิจัยใช้การทดสอบรายคู่ภายหลังด้วย Bonferroni และใช้ Dunnett's T3 ในกรณีที่ความแปรปรวนไม่เป็นเอกพันธ์ และการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ด้วยการวิเคราะห์เนื้อหา

สรุปผลการวิจัย

1. งานวิจัยที่นำมาศึกษาในครั้งนี้มีจำนวน 92 เล่ม เป็นงานวิจัยในมหาวิทยาลัยในกำกับของรัฐ จำนวนมากที่สุด 43 เล่ม (ร้อยละ 46.74) รองลงมาคือ มหาวิทยาลัยราชภัฏ จำนวน 31 เล่ม (ร้อยละ 33.70) มหาวิทยาลัยของรัฐ จำนวน 11 เล่ม (ร้อยละ 11.96) และสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) จำนวน 7 เล่ม (ร้อยละ 7.61) ตามลำดับ ผลการประเมินเพื่อจัดระดับคุณภาพ พบว่า ในภาพรวมส่วนใหญ่เป็นงานวิจัยที่มีคุณภาพอยู่ในระดับดี จำนวน 76 เล่ม (ร้อยละ 82.61) ที่เหลือเป็นงานวิจัยที่มีคุณภาพอยู่ในระดับดีมาก จำนวน 16 เล่ม (ร้อยละ 17.39) คุณภาพงานวิจัยในภาพรวม พบว่า มีคุณภาพอยู่ในระดับดี ($M=2.997$, $SD = 0.247$) เมื่อพิจารณาจำแนกรายด้าน พบว่า ด้านผลการวิเคราะห์ข้อมูล มีค่าเฉลี่ยสูงสุดมีคุณภาพอยู่ในระดับดีมาก ($M=3.875$, $SD = 0.382$) รองลงมาคือ ด้านการสรุป

อภิปรายผลและข้อเสนอแนะ ($M=3.214$, $SD = 0.473$) ด้านวิธีดำเนินการวิจัย ($M= 3.046$, $SD = 0.370$) ด้านการนำเสนอรายงานวิจัย ($M= 2.783$, $SD = 0.340$) ด้านความเป็นมาและความสำคัญของปัญหาวิจัย ($M= 2.634$, $SD = 0.457$) และเอกสารและงานวิจัยที่เกี่ยวข้อง ($M= 2.429$, $SD = 0.340$) ตามลำดับ โดยมีคุณภาพอยู่ในระดับดีทุกด้าน ยกเว้น ด้านการสรุปอภิปรายผลและข้อเสนอแนะที่มีคุณภาพในระดับดีมาก

2. งานวิจัยที่นำมาสังเคราะห์จำนวน 92 เล่ม เมื่อพิจารณาตัวแปรคุณลักษณะของงานวิจัยด้านการพิมพ์และผู้วิจัย พบว่า งานวิจัยมีการตีพิมพ์ในปี 2560 มากที่สุด (ร้อยละ 10.87) ส่วนใหญ่เป็นงานวิจัยจากมหาวิทยาลัยในกำกับของรัฐ (ร้อยละ 46.74) โดยมีจุฬาลงกรณ์มหาวิทยาลัย และมหาวิทยาลัยราชภัฏมหาสารคาม เป็นสถาบันที่ผลิตงานวิจัยมากที่สุด (ร้อยละ 15.22) ซึ่งเป็นงานวิจัยที่ผลิตโดยกลุ่มสาขาการวัดและประเมินผล และกลุ่มสาขาการวิจัยและสถิติ (ร้อยละ 45.65) ส่วนใหญ่เป็นงานวิจัยระดับปริญญาโท (ร้อยละ 86.96) และผลิตโดยนักวิจัยเพศหญิง (ร้อยละ 82.61) สำหรับตัวแปรคุณลักษณะของงานวิจัยด้านเนื้อหาสาระงานวิจัย พบว่า งานวิจัยมีวัตถุประสงค์เพื่อสร้าง/พัฒนาเครื่องมือวัดมากที่สุด (ร้อยละ 33.46) โดยงานวิจัยมีการพัฒนาเครื่องมือวัดคุณลักษณะอันพึงประสงค์มากที่สุด (ร้อยละ 13.04) งานวิจัยส่วนใหญ่ใช้แนวคิด/ทฤษฎีหลักที่เกิดจากการสังเคราะห์ (ร้อยละ 31.52) ทั้งนี้ เครื่องมือวัดในองค์ประกอบการสร้างสรรค์และนวัตกรรมมากที่สุด (ร้อยละ 30.43) ซึ่งเกินครึ่งหนึ่งเป็นเครื่องมือวัดที่ไม่อิงเนื้อหาวิชา (content free) (ร้อยละ 60.87) มีการศึกษาในภูมิภาคตะวันออกเฉียงเหนือมากที่สุด (ร้อยละ 38.04) โดยส่วนใหญ่เป็นตัวอย่างจากสังกัดสำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) มากที่สุด (ร้อยละ 82.61) มีการศึกษากับนักเรียนชั้นมัธยมศึกษาตอนต้นมากที่สุด (ร้อยละ 25.00) งานวิจัยส่วนใหญ่มีการศึกษาทั้งวิธีการสร้างและการวิเคราะห์คุณภาพมากที่สุด (ร้อยละ 70.65) เครื่องมือวัดที่สร้างขึ้นเกือบทั้งหมดผู้วิจัยเป็นผู้พัฒนาเครื่องมือวัดขึ้นเอง (ร้อยละ 98.91) งานวิจัยส่วนใหญ่มีและนำเสนอนิยามเชิงปฏิบัติการ (ร้อยละ 82.61) รวมถึงมีการสร้างตารางกำหนดโครงสร้างของเครื่องมือวัดและนำเสนอไว้ในงานวิจัย (ร้อยละ 70.65) อย่างไรก็ตามงานวิจัยทุกเล่มมีการระบุขั้นตอนในการพัฒนาเครื่องมือวัด (ร้อยละ 100) งานวิจัยมีการพัฒนาเครื่องมือวัดเป็นแบบสอบอัตโนมัติมากที่สุด (ร้อยละ 29.35) เครื่องมือพัฒนาขึ้นมีลักษณะ 4 ตัวเลือก/เกณฑ์ มากที่สุด (ร้อยละ 30.43) ซึ่งมีการทดลองใช้เครื่องมือจำนวน 3 ครั้ง มากที่สุด (ร้อยละ 44.57) มีการวิเคราะห์หาคุณภาพเครื่องมือวัดด้วยทฤษฎีการทดสอบแบบดั้งเดิม (CTT) มากที่สุด (ร้อยละ 84.78) โดยครึ่งหนึ่งของเครื่องมือวัดมีการตรวจให้คะแนนมากกว่าสองค่า (ร้อยละ 56.52) ทั้งนี้ งานวิจัยมีการสร้างเกณฑ์ปกติระดับท้องถิ่นมากที่สุด (ร้อยละ 35.87) โดยครึ่งหนึ่งเป็นการสร้างเกณฑ์ปกติด้วยวิธีคะแนนมาตรฐานที่ปกติ (ร้อยละ 58.70) และงานวิจัยส่วนใหญ่มีการสร้างคู่มือการใช้เครื่องมือวัด (ร้อยละ 76.09) สำหรับตัวแปรคุณลักษณะของงานวิจัยด้านวิธีวิทยาการวิจัย พบว่า งานวิจัยมีการใช้สูตรของ Yamane (1973) ในการกำหนดขนาดของตัวอย่างมากที่สุด (ร้อยละ 21.74) ทั้งนี้ มีงานวิจัยจำนวนครึ่งหนึ่งใช้สุ่มแบบหลายขั้นตอน (ร้อยละ 50.00) และพบว่า งานวิจัยมีการวิเคราะห์ค่าความเที่ยง และค่าความตรงทุกเล่ม (ร้อยละ 100) ในส่วนของตัวแปรคุณลักษณะของงานวิจัยด้านคุณภาพงานวิจัย พบว่า งานวิจัยส่วนใหญ่มีคุณภาพอยู่ในระดับดี (ร้อยละ 82.61)

3. ผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา พบว่า

3.1 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความเที่ยง มีจำนวน 11 ตัวแปร ได้แก่ ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, เพศของผู้วิจัย, ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนตัวเลือก/เกณฑ์, การตรวจให้คะแนน, จำนวนข้อคำถามในแต่ละฉบับ และวิธีการเลือกตัวอย่าง

3.2 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความตรง มีเพียงตัวแปรเดียวคือ เนื้อหาที่เกี่ยวข้อง

3.3 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานความยาก มีจำนวน 6 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทงานวิจัย, การสร้างเครื่องมือวัด, ประเภทของเครื่องมือวัด จำนวนตัวเลือก/เกณฑ์ และวิธีการเลือกตัวอย่าง

3.4 คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานอำนาจจำแนก มีจำนวน 10 ตัวแปร ได้แก่ ปีที่พิมพ์งานวิจัย, ประเภทสถาบันที่ผลิตงานวิจัย, สถาบันที่ผลิตงานวิจัย, สาขาที่ผลิตงานวิจัย, เพศของผู้วิจัย, ภูมิภาค, ระดับชั้นของตัวอย่าง, ประเภทของเครื่องมือวัด, จำนวนข้อคำถามในแต่ละฉบับ และจำนวนผู้สอบในแต่ละฉบับ

4. ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา พบว่า

4.1 การสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา พบว่า ในภาพรวมการพัฒนาเครื่องมือ ในงานวิจัยแต่ละเล่มมีความคล้ายคลึงกัน โดยผู้วิจัยแบ่งการพัฒนาเครื่องมือ ออกเป็น 3 ขั้นตอนหลัก ดังนี้

4.1.1 ขั้นตอนการสร้างข้อคำถาม ประกอบด้วย 1) กำหนดวัตถุประสงค์ของการสร้างเครื่องมือ 2) ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง 3) กำหนดองค์ประกอบและตัวบ่งชี้ 4) สร้างตารางกำหนดโครงสร้างของเครื่องมือ 5) สํารวจ สัมภาษณ์ความคิดเห็นกลุ่มเป้าหมายเพื่อนำสถานการณ์ไปสร้างข้อคำถาม และ 6) สร้างข้อคำถาม และเกณฑ์การให้คะแนน

4.1.2 ขั้นตอนการหาคุณภาพ ประกอบด้วย

1) หาความตรงตามเนื้อหาโดยผู้เชี่ยวชาญ ซึ่งสามารถพิจารณาจากการวิเคราะห์ความสอดคล้องระหว่างข้อคำถามและนิยามศัพท์หรือวัตถุประสงค์ในการวัด (Item Objectice Congruence: IOC) หรือ การพิจารณาความเหมาะสมของข้อคำถามด้วยแบบประเมินแบบมาตรประมาณค่า

2) ทดลองใช้เครื่องมือ (try out) เพื่อตรวจสอบคุณภาพเครื่องมือ ในด้านความตรง ความเที่ยง ความยาก อำนาจจำแนก รวมถึงการตรวจสอบด้านพฤติกรรม การใช้ภาษาและเวลา การวิเคราะห์การทำหน้าที่ต่างกันของข้อคำถาม และการวิเคราะห์ความไม่แปรเปลี่ยน โดยสามารถเลือกวิธีวิเคราะห์ได้ตามความเหมาะสม ตามแนวคิดทฤษฎีการทดสอบแบบดั้งเดิม (Classical Test Theory: CTT) และ/หรือทฤษฎีการทดสอบแนวใหม่ (modern test theory เช่น IRT และ G theory) เป็นต้น

4.1.3 ขั้นตอนการสร้างปกติวิสัยและคู่มือการใช้เครื่องมือ ขั้นตอนการสร้างปกติวิสัย (norm) เป็นกระบวนการสุดท้ายของการพัฒนาเครื่องมือวัดสมรรถนะ เพื่อใช้เป็นเกณฑ์ในการแปลผลระดับสมรรถนะของนักเรียน ซึ่งมีหลายวิธี เช่น วิธีเปอร์เซ็นต์ไทล์ วิธีคะแนนมาตรฐานที่ วิธีคะแนนมาตรฐานที่ปกติ และวิธีสเตโน

4.2 สรุปเครื่องมือที่นิยมใช้ในแต่ละสมรรถนะ แบ่งออกเป็น 4 กลุ่ม ดังนี้

กลุ่มที่ 1 สมรรถนะที่ใช้เครื่องมือประเภทมาตรประมาณค่า ได้แก่ จิตสำนึกต่อโลก ภาวะผู้นำ การรู้เท่าทันสื่อและสารสนเทศ และการรู้เท่าทันเทคโนโลยี

กลุ่มที่ 2 สมรรถนะที่ใช้เครื่องมือประเภทแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า ได้แก่ ความรู้เรื่องพลเมือง การแก้ปัญหา ทักษะด้านการทำงานเป็นทีม ความร่วมมือ และความรับผิดชอบ

กลุ่มที่ 3 สมรรถนะที่ใช้เครื่องมือประเภทแบบสอบหลายตัวเลือก ให้คะแนน 2 ค่า ได้แก่ การคิดอย่างมีวิจารณ์ญาณ และการสื่อสาร

กลุ่มที่ 4 สมรรถนะที่ใช้เครื่องมือประเภทแบบสอบอัตนัย ได้แก่ การสร้างสรรค์และนวัตกรรม

อภิปรายผลการวิจัย

จากผลการวิจัย ผู้วิจัยนำเสนอการอภิปรายผลการวิจัย ออกเป็น 3 ประเด็น ประกอบด้วย 1) ผลการประเมินคุณภาพงานวิจัย 2) คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา และ 3) องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

1. ผลการประเมินคุณภาพงานวิจัย พบว่างานวิจัยมีคุณภาพอยู่ในระดับดี ร้อยละ 82.61 และดีมาก ร้อยละ 17.39 ทั้งนี้เป็นเพราะงานวิจัยที่นำมาใช้ในการวิจัยครั้งนี้เป็นงานวิจัยประเภทวิทยานิพนธ์ในระดับบัณฑิตศึกษาทั้งระดับมหาบัณฑิต และดุษฎีบัณฑิต รวมถึงเป็นรายงานการวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) การทำงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัศึกษามีคุณภาพอยู่ในระดับดีและดีมากนั้น เป็นเพราะงานวิจัยประเภทวิทยานิพนธ์ อยู่ภายใต้การดูแล ให้ข้อเสนอแนะ และควบคุมคุณภาพโดยอาจารย์ที่ปรึกษาวิทยานิพนธ์ รวมถึงผ่านกระบวนการสอบจากคณะกรรมการสอบวิทยานิพนธ์มีจำนวน 3-5 คน ซึ่งเป็นผู้เชี่ยวชาญในการพัฒนาเครื่องมือวัดจึงทำให้ได้ผลงานที่มีคุณภาพในระดับหนึ่ง ในขณะที่รายงานการวิจัยที่ได้รับทุนสนับสนุนการวิจัยจากสถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) นักวิจัยผู้ได้รับทุนต่างเป็นคณาจารย์ในสาขาวิชาการวัดและประเมินผล รวมถึงวิจัยและสถิติทางการศึกษา ซึ่งเป็นผู้ที่มีความรู้ และความเชี่ยวชาญในการพัฒนาเครื่องมือวัด ทำให้งานวิจัยในภาพรวมมีคุณภาพในระดับดีขึ้นไป ทั้งนี้ผลการวิจัยดังกล่าวสอดคล้องกับงานวิจัยของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ จำนวน 75 เล่ม ซึ่งมีการตีพิมพ์ระหว่างปี พ.ศ. 2523-2551 พบว่า ผลการประเมินคุณภาพงานวิจัยอยู่ในระดับดีมาก มีจำนวนมากที่สุด จำนวน 45 เล่ม (ร้อยละ 60.00) รองลงมาคือ มีคุณภาพในระดับดี จำนวน 29 เล่ม (ร้อยละ 38.70) และคุณภาพในระดับปานกลางมีเพียงจำนวน 1 เล่ม (ร้อยละ 1.30) รวมถึงงานวิจัยของ รัตนะ บัวสนธ์ และคณะ (2557) ที่ได้ประเมินคุณภาพงานวิจัย ประเภทวิทยานิพนธ์ของนิสิตระดับดุษฎีบัณฑิต มหาวิทยาลัยนเรศวร ระหว่างปี พ.ศ. 2550-2554 จำนวน 30 เรื่อง พบว่า ผลการประเมินคุณภาพงานวิจัยร้อยละ 53.33 อยู่ในระดับดีมาก และร้อยละ 43.33 มีคุณภาพระดับดี

เมื่อพิจารณาผลการประเมินคุณภาพรายด้าน ในด้านความเป็นมาและความสำคัญของปัญหาวิจัยมีผลการประเมินในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมากมี 1 รายการ คือ ชื่อเรื่องมีความชัดเจน น่าสนใจ กล่าวคือ ชื่อเรื่องงานวิจัยโดยส่วนใหญ่มีการระบุถึงปัญหาวิจัย ระบุตัวแปรสำคัญที่ศึกษาครบถ้วน และระบุตัวอย่างที่ศึกษาได้อย่างครบถ้วน อย่างไรก็ตาม มีบางประเด็นที่ควรให้ความสำคัญมากขึ้น ได้แก่ เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผลที่มีระดับคุณภาพค่อนข้างต่ำ และการนิยามศัพท์เฉพาะมีความชัดเจน ที่มีระดับคุณภาพปานกลาง ทั้งนี้ เนื่องจากงานวิจัยที่พัฒนาเครื่องมือวัดถึงแม้จะเป็นเรื่องที่น่าสนใจ แต่ผู้วิจัยควรให้เหตุผลประกอบว่าเป็นเรื่องที่ต้องรู้ความรู้อย่างไร เป็นประเด็นที่ไม่เคยมีผู้ใดศึกษามาก่อน และเป็นปัญหาเร่งด่วนที่ต้องการพัฒนาเครื่องมือวัดเนื่องจากเหตุผลใด สำหรับการนิยามศัพท์เฉพาะมีความชัดเจน ที่มีผลการประเมินในระดับปานกลางนั้น เนื่องมาจากงานวิจัยส่วนใหญ่มีการนิยามเชิงทฤษฎี หรือนิยามเชิงโครงสร้าง (construct definition) ที่อิงทฤษฎีได้อย่างครบถ้วน ซึ่งนิยามเชิงทฤษฎี

นั้น สุวิมล ติรภานันท์ (2551) กล่าวว่า เป็นการให้ความหมายเชิงทฤษฎี และแนวคิดต่าง ๆ ที่เกี่ยวข้อง ทำให้เข้าใจถึงลักษณะและความหมายของตัวแปรมากขึ้น แต่ยังไม่สามารถสังเกตได้ หรือยังไม่เห็นส่วนประกอบอย่างชัดเจน มีลักษณะเป็นนามธรรม แต่ทั้งนี้ผลการวิจัยพบว่า ยังขาดการนิยามเชิงปฏิบัติการ (operational definition) ซึ่งเป็นการให้ความหมายเกี่ยวกับตัวแปรที่ศึกษาให้อยู่ในรูปของพฤติกรรมหรือรูปรธรรมซึ่งสามารถวัดได้ รวมทั้งต้องมีเนื้อหาสอดคล้องกับเนื้อหาในนิยามเชิงทฤษฎี (คณะกรรมการสภาวิจัยแห่งชาติ สาขาศึกษา, 2555; บุญใจ ศรีสถิตยัณรากร, 2555) อย่างไรก็ตาม บุญใจ ศรีสถิตยัณรากร (2555) กล่าวว่า ตัวแปรที่ต้องมีการนิยามเชิงปฏิบัติการ คือ ตัวแปรที่มีนิยามเชิงทฤษฎีที่เป็นนามธรรม (abstract concept) บุคคลทั่วไปหรือกลุ่มตัวอย่างรับรู้ความหมายของตัวแปรหรือแปลความหมายตัวแปรไม่ตรงกัน ทั้งนี้หากขาดการนิยามเชิงปฏิบัติการ หรือมีการเขียนนิยามเชิงปฏิบัติการที่ไม่กระจ่างชัดเจน จะทำให้เนื้อหาในนิยามเชิงปฏิบัติการไม่ถูกต้อง ย่อมส่งผลให้เครื่องมือวัดที่สร้างขึ้นขาดความตรงตามเนื้อหา อย่างไรก็ตามผลการประเมินดังกล่าวมีความสอดคล้องกับผลการวิจัยของ อาภาพรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด จำนวน 61 เล่มที่มีการตีพิมพ์เผยแพร่ระหว่างปี พ.ศ.2509-2551 และผลการวิจัยของ สรัญญา จันทรชูสกุล (2561) ที่ได้ประเมินคุณภาพงานวิจัยนวัตกรรมการพัฒนาความคิดสร้างสรรค์ของนักเรียนระดับประถมศึกษาจำนวน 118 เล่ม ที่มีการตีพิมพ์เผยแพร่ระหว่างปี พ.ศ. 2546-2558 ผลการวิจัยพบว่า เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล มีระดับคุณภาพค่อนข้างต่ำ และการนิยามศัพท์เฉพาะมีความชัดเจน มีระดับคุณภาพปานกลาง เช่นเดียวกัน

สำหรับผลคุณภาพด้านเอกสารและงานวิจัยที่เกี่ยวข้องในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก มี 1 รายการ คือ เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ อย่างไรก็ตามมีบางประเด็นที่ควรให้ความสำคัญมากขึ้น ได้แก่ เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย ที่มีผลการประเมินอยู่ในระดับค่อนข้างต่ำ รวมถึงเอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย และเอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศมีสัดส่วนที่เหมาะสม ที่มีผลการประเมินอยู่ในระดับปานกลาง สำหรับประเด็นของความทันสมัยของเอกสารและงานวิจัยนั้น พบว่า งานวิจัยโดยส่วนใหญ่มีการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณไม่ถึงร้อยละ 30 นอกจากนี้ เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทยเป็นส่วนมาก มีสัดส่วนเอกสารและงานวิจัยภาษาต่างประเทศไม่ถึงร้อยละ 20 ทั้งนี้ในการทบทวนเอกสารหรืองานวิจัยที่เกี่ยวข้อง นักวิจัยอาจจะอ้างอิงเอกสารต้นฉบับที่เป็นเอกสารชั้นครูได้ แต่ควรมีเอกสารที่ทันสมัยที่แสดงให้เห็นพัฒนาการจากแนวคิดนั้นด้วย โดยเฉพาะอย่างยิ่งงานวิจัยที่นำมาใช้ศึกษาอ้างอิงควรเป็นงานวิจัยใหม่ ๆ ที่ทันสมัย เนื่องจากหากเป็นการศึกษาจากงานวิจัยเก่า ๆ การนำไปใช้ประโยชน์ของนักวิจัยจะอยู่ในวงที่จำกัด เพราะงานวิจัยเป็นการทำซ้ำ ทำให้สูญเสียทรัพยากรโดยไม่จำเป็น (คณะกรรมการสภาวิจัยแห่งชาติ สาขาศึกษา, 2555) ผลการประเมินดังกล่าวมีความสอดคล้องกับผลการวิจัยของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ และงานวิจัยของ อาภาพรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด พบว่า เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย มีผลการประเมินอยู่ในระดับค่อนข้างต่ำเช่นเดียวกัน

ผลการประเมินคุณภาพงานวิจัยด้านวิธีดำเนินการวิจัยในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก มี 5 รายการ ได้แก่ การออกแบบการวิจัยสอดคล้องกับปัญหาการวิจัย, กำหนดประชากรและตัวอย่างได้อย่างเหมาะสมกับการวิจัย, ขนาดตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม, วิธีการตรวจสอบคุณภาพเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม และตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม อย่างไรก็ตามมีบางประเด็นที่ควรให้ความสำคัญมากขึ้น ได้แก่ การกำหนดจุดมุ่งหมายในการวัดสมรรถนะใน

ศตวรรษที่ 21 มีความชัดเจน และการสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา ซึ่งมีผลการประเมินที่อยู่ในระดับปานกลาง กล่าวคือ งานวิจัยบางส่วนยังไม่มีจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 หรือมีการกำหนดแต่ไม่ชัดเจน หรือไม่สอดคล้องกับการนำผลไปใช้ประโยชน์ อย่างไรก็ตาม Newton (2007 อ้างถึงใน ญัฐภรณ์ หลาวทอง, 2559) กล่าวว่า การกำหนดจุดมุ่งหมายของการสร้างเครื่องมือวัดทางการวิจัย และการวัดประเมินผลทางการศึกษา แบ่งออกเป็น 18 ประการ ได้แก่ 1) ใช้เพื่อการประเมินทางสังคม 2) ใช้เพื่อการพัฒนา 3) ใช้เพื่อการติดตามกำกับผู้เรียน 4) ใช้เพื่อการเทียบโอน 5) ใช้เพื่อการจัดวางตำแหน่ง 6) ใช้เพื่อการวินิจฉัย 7) ใช้เพื่อการแนะแนว 8) ใช้เพื่อวัดคุณสมบัติ 9) ใช้เพื่อการคัดเลือก 10) ใช้เพื่อกำหนดใบอนุญาต 11) ใช้เพื่อการเลือกโรงเรียน 12) ใช้เพื่อการติดตามสถาบัน 13) ใช้เพื่อการจัดสรรทรัพยากร 14) ใช้เพื่อดำเนินการขององค์กร 15) ใช้เพื่อประเมินโปรแกรม 16) ใช้เพื่อการติดตามระบบ 17) ใช้เพื่อการเปรียบเทียบ/เทียบเคียง และ 18) ใช้เพื่อการประกันคุณภาพระดับชาติ สำหรับการสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหาที่มีผลการประเมินที่อยู่ในระดับปานกลางนั้น พบว่า มีงานวิจัย ร้อยละ 18.48 ไม่กำหนดแผนผังการสร้างข้อคำถาม ในขณะที่งานวิจัยร้อยละ 10.87 มีการกำหนดแผนผังการสร้างข้อคำถาม แต่ไม่นำเสนอในเล่มรายงานวิจัย อย่างไรก็ตาม มีงานวิจัยร้อยละ 70.65 มีการกำหนดแผนผังการสร้างข้อคำถาม แต่ยังมีบางส่วนที่ขาดความเหมาะสม เช่น ไม่มีการแยกรายละเอียดตามองค์ประกอบ หรือตัวชี้วัดที่ต้องการวัด หรือขาดการระบุเลขข้อ จำนวนข้อที่วัดตามองค์ประกอบนั้น ๆ เป็นต้น ผลการประเมินดังกล่าวมีความสอดคล้องกับผลการวิจัยของ งานวิจัยของ อากาศพรหม ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด พบว่า การกำหนดจุดมุ่งหมายในการสร้างแบบวัดมีความชัดเจนอยู่ในระดับปานกลาง แต่การสร้างแบบวัดมีการสร้างตารางกำหนดแผนผังการสร้างข้อสอบที่ครอบคลุมเนื้อหา มีคุณภาพในระดับต่ำ

ผลการประเมินคุณภาพงานวิจัยด้านผลการวิเคราะห์ข้อมูล ในภาพรวมอยู่ในระดับดีมาก เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก ทั้ง 2 รายการ ได้แก่ ลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูล และการแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องชัดเจน กล่าวคือ มีการนำเสนอผลการวิเคราะห์ข้อมูลที่ควรจะนำเสนออย่างครบถ้วน วิธีนำเสนอเหมาะสม เช่น นำเสนอในรูปของตาราง มีการอธิบายผลการวิเคราะห์ รวมถึงมีการแปลความหมายหรือสรุปผลการวิเคราะห์ข้อมูลครบทุกส่วน ถูกต้องและสอดคล้องกับผลวิเคราะห์ที่ได้ ทั้งนี้ วรรณิ แกมเกตุ (2555) กล่าวว่า การรายงานผลการวิเคราะห์ข้อมูลควรประกอบด้วย 3 ส่วนหลัก คือ 1) การวิเคราะห์ข้อมูลเบื้องต้นเกี่ยวกับกลุ่มตัวอย่างที่ใช้ในการวิจัย 2) ผลการวิเคราะห์ข้อมูลเพื่อตรวจสอบว่าข้อมูลเป็นไปตามข้อตกลงเบื้องต้นหรือไม่ และ 3) ผลการวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย โดยการเขียนบรรยายผลการวิเคราะห์ข้อมูลควรเขียนในลักษณะการแปลความหรือตีความค่าสถิติจากผลการวิเคราะห์ มิใช่เป็นการอ่านตาราง และในกรณีที่เป็นสถิติใหม่ ควรเขียนอธิบายความหมายของค่าสถิตินั้นก่อน แล้วนำไปสู่การตีความ นอกจากนี้ ผู้วิจัยควรออกแบบการนำเสนอตารางที่ต้องการนำเสนอในเนื้อหาในลักษณะเดียวกัน ควรออกแบบในตารางเดียวกัน เพื่อให้การนำเสนอมีความกระชับ และควรใช้รูปแบบตารางเดียวกันตลอดทั้งเล่ม ผลการวิจัยดังกล่าว สอดคล้องกับผลการศึกษาของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์ ที่พบว่ามีผลการประเมินทั้ง 2 รายการอยู่ในระดับดีมากเช่นกัน แต่ผลการศึกษาของ อากาศพรหม ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด พบว่า งานวิจัยมีลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูลอยู่ในระดับดีมาก แต่มีการแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องอยู่ในระดับดี

ผลการประเมินคุณภาพงานวิจัยด้านการสรุปอภิปรายผลและข้อเสนอแนะ ในภาพรวมอยู่ในระดับดีมาก เมื่อพิจารณาในรายละเอียด พบว่า ผลการประเมินที่อยู่ในระดับดีมาก 2 รายการ คือ สรุปผลการวิจัยอย่างถูกต้อง และมีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีต กล่าวคือ งานวิจัยส่วนใหญ่ สามารถสรุปผลวิจัยถูกต้องสอดคล้องกับวัตถุประสงค์การวิจัย ตอบปัญหาวิจัยครอบคลุมทุกหัวข้อ ใช้ข้อความกระชับ อ่านเข้าใจง่าย อ่านแล้วเข้าใจถึงประเด็นที่ต้องการสรุปในหัวข้อนั้น ๆ อย่างชัดเจน รวมถึงมีการอภิปรายผลการวิจัยที่ได้กับผลงานวิจัยในอดีตอย่างสมเหตุสมผล ครอบคลุมทุกประเด็นปัญหาวิจัย และมีผลการประเมินที่อยู่ในระดับดี 1 รายการ คือ ข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์ กล่าวคือ มีการให้ข้อเสนอแนะถึงการนำผลการวิจัยไปใช้มาจากผลการวิจัย มีเหตุผลรองรับเพียงพอ แต่อย่างไรก็ดี สุวิมล ติรภานันท์ (2557) กล่าวว่าสิ่งที่ผู้วิจัยจำเป็นต้องคำนึงเสมอในการเขียนข้อเสนอแนะ คือ ข้อเสนอแนะนั้นจะต้องเป็นสาระที่เกิดขึ้นจากผลการวิจัย มิใช่ความรู้สึก หรือความคิดเห็นของผู้วิจัย โดยจะต้องนำเสนอในสิ่งใหม่ ไม่ใช่เรื่องทราบโดยทั่วไป เป็นข้อเสนอที่สามารถปฏิบัติได้ ขอบเขตเงื่อนไขที่กำหนดไว้ โดยจะต้องเป็นข้อเสนอที่ต่อเนื่องหรือต่อยอดจากงานวิจัย ผลการวิจัยดังกล่าวสอดคล้องกับผลการศึกษาของ สุวิมล เสวกสุริยวงศ์ (2553) ที่ได้ประเมินคุณภาพงานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจัยทางคณิตศาสตร์ และสร้อยัญญา จันทร์ชูสกุล (2561) ที่ได้ประเมินคุณภาพงานวิจัยนวัตกรรม การพัฒนาความคิดสร้างสรรค์ของนักเรียนระดับประถมศึกษา พบว่า การสรุปผลการวิจัยอย่างถูกต้องมีคุณภาพอยู่ในระดับดีมาก และข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์อยู่ในระดับดีเช่นเดียวกัน ทั้งนี้ ยกเว้นรายการที่มีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีตที่ผลการประเมินอยู่ในระดับดี

ผลการประเมินคุณภาพงานวิจัยด้านการนำเสนอรายงานวิจัย ในภาพรวมอยู่ในระดับดี เมื่อพิจารณาในรายละเอียดพบว่า ผลการประเมินที่อยู่ในระดับดีมาก 1 รายการ คือ รูปแบบรายงานถูกต้องตามหลักวิชา อย่างไรก็ตาม มีประเด็นที่ควรให้ความสำคัญมากขึ้น คือ เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ ที่มีระดับคุณภาพค่อนข้างต่ำ กล่าวคือ งานวิจัยส่วนใหญ่มีประโยชน์ในด้านวิชาการระดับการประยุกต์ทฤษฎีเพื่อนำไปใช้ เนื่องจากงานวิจัยมีจุดมุ่งหมายหลักในการสร้างและตรวจสอบคุณภาพของเครื่องมือวัดที่สร้างขึ้น จากแนวคิดหรือทฤษฎีที่มีนักวิชาการเสนอไว้แล้ว หรือมีการสังเคราะห์จากหลาย ๆ แนวคิดหรือทฤษฎี เพื่อใช้ป้อนองค์ประกอบหรือตัวบ่งชี้ในการสร้างเครื่องมือ แต่ยังไม่ถึงขั้นได้องค์ความรู้ใหม่ หรือปรับปรุงแนวคิดและทฤษฎี จึงทำให้ผลการประเมินมีระดับคุณภาพค่อนข้างต่ำ ซึ่งผลการประเมินดังกล่าวสอดคล้องกับงานวิจัยของ อาภาพรรณ ประทุมไทย (2553) ที่ได้ประเมินคุณภาพงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด ซึ่งมีผลการประเมินไปในทิศทางเดียวกันว่า งานวิจัยที่ศึกษาในบริบทของการสร้างแบบวัดมีระดับคุณภาพของการนำไปใช้ประโยชน์ได้ในทางวิชาการได้ค่อนข้างต่ำ

2. คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

ในการวิจัยครั้งนี้ผู้วิจัยได้ใช้การวิเคราะห์ห่อภิมาณ ซึ่งเป็นการสังเคราะห์เชิงปริมาณงานวิจัยรูปแบบหนึ่งมาใช้ในการวิเคราะห์ข้อมูล โดยตัวแปรตามในการวิจัยเป็น ค่าดัชนีมาตรฐานคุณภาพจากเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา มี 4 ประเภท คือ ค่าดัชนีมาตรฐานความเที่ยง ค่าดัชนีมาตรฐานความตรง ค่าดัชนีมาตรฐานความยาก และค่าดัชนีมาตรฐานอำนาจจำแนก โดยงานที่นำมาใช้ในการวิเคราะห์ห่อภิมาณมาจากงานวิจัยจำนวน 90 เล่ม ที่มีการตรวจสอบคุณภาพเครื่องมือวัดตามแนวคิดของทฤษฎีทดสอบแบบดั้งเดิม โดยค่าดัชนีมาตรฐานทั้ง 4 ประเภท มีความสำคัญดังนี้

ค่าดัชนีมาตรฐานความเที่ยง Miller, et al (2009) ได้กล่าวถึงธรรมชาติของความเที่ยงไว้ว่า ความเที่ยงของผลคะแนนที่ได้จากการทดสอบหรือผลการประเมินอื่น ๆ ที่ได้จากการใช้เครื่องมือต่าง ๆ เพื่อวัดและประเมิน ซึ่งไม่ได้หมายถึงความเที่ยงของเครื่องมืออื่น ๆ ด้วยเหตุนี้ค่าความเที่ยงอาจจะมีค่าที่แตกต่างกันได้ขึ้นอยู่กับกลุ่มและสถานการณ์ที่ได้นำเครื่องมืออื่น ๆ ไปใช้ ดังนั้น จึงเป็นความเที่ยงของผลการวัดหรือผลการประเมินมากกว่าจะกล่าวว่าเป็นความเที่ยงของแบบสอบหรือแบบประเมิน ทั้งนี้ การประมาณค่าความเที่ยงจะนำเสนอด้วยค่าสถิติที่เรียกว่า สัมประสิทธิ์ความเที่ยง (reliability coefficient) ที่ได้จากวิธีการประมาณค่าความเที่ยงหลายวิธี ซึ่งค่าความเที่ยงที่ประมาณค่าได้นั้นมักขึ้นอยู่กับประเภทของวิธีการประเมินค่า ด้วยเหตุนี้ผู้วิจัยจึงต้องพิจารณาเลือกวิธีการประมาณค่าความเที่ยงให้เหมาะสมกับจุดมุ่งหมาย นอกจากนี้ ความเที่ยงเป็นสิ่งที่จำเป็นแต่ไม่เพียงพอที่เป็นเงื่อนไขสำหรับความตรงกล่าวคือ ค่าความเที่ยงต่ำบ่งบอกถึงความตรงที่ต่ำ แต่ค่าความเที่ยงที่สูงค่าความตรงอาจจะสูงหรือต่ำก็ได้ ในการวิจัยครั้งนี้ พบว่า มีค่าความเที่ยงของเครื่องมือวัดจำนวน 216 ค่า มีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยง (M) เท่ากับ .778 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .163 เมื่อพิจารณาตามเกณฑ์ที่เสนอโดย Isaacs, Zara, & Herbert (2013) พบว่า ค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงอยู่ในระดับสูง

ค่าดัชนีมาตรฐานความตรง ศิริชัย กาญจนวาสี, 2556 และ Miller, et al, 2009 ได้กล่าวถึงธรรมชาติของความตรง ความตรงเป็นการใช้ผลและการแปลความหมายของผลที่ได้กับกลุ่มใดกลุ่มหนึ่งได้อย่างเหมาะสม ไม่ได้เป็นความตรงของแบบสอบหรือความตรงของกระบวนการประเมินโดยตรง ความตรงเป็นระดับความเข้ม ไม่ใช่แต่เพียงมีความตรงหรือไม่มีความตรงเท่านั้น แต่เป็นระดับที่สามารถระบุได้ ความตรงเป็นการใช้ผลหรือแปลผลในเรื่องใดเรื่องหนึ่งเป็นการเฉพาะตามจุดประสงค์ที่กำหนด ความตรงเป็นแนวคิดเอกลักษณ์ (unitary concept) กล่าวคือ ความตรงเป็นคุณลักษณะเดียวที่จำเป็นต้องตรวจสอบด้วยวิธีการหลายวิธี เพื่อให้ได้หลักฐานต่าง ๆ สำหรับสนับสนุนหรือยืนยันความตรง ทั้งนี้ในการประมาณค่าความตรงสามารถจำแนกได้เป็น 3 ประเภทใหญ่ ๆ ได้แก่ (1) ความตรงตามเนื้อหา (content validity) (2) ความตรงตามเกณฑ์สัมพันธ์ (criterion-related validity) แบ่งเป็น ความตรงตามสภาพ/ร่วมสมัย (concurrent validity) และความตรงเชิงทำนาย (predictive validity) และ (3) ความตรงเชิงโครงสร้าง หรือความตรงเชิงทฤษฎี (construct validity) อย่างไรก็ตาม ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดขอบเขตในการวิเคราะห์หรืออภิปรายเฉพาะความตรงตามเกณฑ์สัมพันธ์ แต่จากผลการศึกษาพบว่า มีเพียงความตรงตามสภาพ จำนวน 39 ค่า มีค่าเฉลี่ยดัชนีมาตรฐานความตรง (M) เท่ากับ .596 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .199

ค่าดัชนีมาตรฐานความยาก เป็นการวิเคราะห์คุณภาพของเครื่องมือวัดประเภทแบบสอบ โดยเป็นดัชนีของคุณภาพข้อสอบรายข้อ กล่าวคือ เป็นร้อยละหรือสัดส่วนของผู้สอบที่ตอบข้อสอบข้อนั้นได้ถูกต้อง มีค่าอยู่ระหว่าง 0 ถึง 1 กล่าวคือ หากค่าเข้าใกล้ 0 แสดงว่าข้อสอบข้อนั้นยาก แต่หากค่าเข้าใกล้ 1 แสดงว่าข้อสอบข้อนั้นง่าย ทั้งนี้ Reynolds, et al. (2009) กล่าวว่า ข้อสอบควรมีค่าความยากของข้อสอบ (P) อยู่ระหว่าง 0.2-0.8 จึงเป็นข้อสอบที่มีความยากง่ายพอเหมาะ และค่าเฉลี่ยความยากของข้อสอบทั้งหมดควรมีค่าประมาณ 0.5 เนื่องจากเป็นค่าที่ทำให้แบบสอบทั้งหมดมีค่าความเที่ยงและความตรงสูงที่สุดใน การวิจัยครั้งนี้ ผู้วิจัยนำค่าความยากของข้อสอบแต่ละข้อมาหาเฉลี่ยเป็นค่าดัชนีมาตรฐานความยาก เครื่องมือวัดทั้งหมด พบว่า มีจำนวน 94 ค่า มีค่าเฉลี่ยดัชนีมาตรฐานความยาก (M) เท่ากับ .537 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .113

ค่าดัชนีมาตรฐานอำนาจจำแนก เป็นความสามารถของข้อสอบในการจำแนกหรือแยกให้เห็นความแตกต่างระหว่างผู้สอบที่มีผลสัมฤทธิ์ทางการเรียนแตกต่างกันออกจากกันได้ โดยเป็นดัชนีของคุณภาพข้อสอบรายข้อ มีค่าอยู่ระหว่าง -1 ถึง 1 แต่ค่าอำนาจจำแนกที่ดีต้องมีค่าเป็นบวก ทั้งนี้ Reynolds,

et.al. (2009) เสนอว่า ค่าอำนาจจำแนกที่ดีควรมีค่าตั้งแต่ 0.3 ขึ้นไป หากข้อสอบข้อใดที่มีค่าอำนาจจำแนกต่ำกว่า 0.3 จะต้องพิจารณาว่าควรมีการปรับปรุงหรือตัดข้อสอบข้อนั้นออกไป อย่างไรก็ตามสำหรับนักวิชาการทางวัดประเมินผลของไทย กล่าวว่า ค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไปเป็นค่าอำนาจจำแนกที่พอใช้ได้ (โชติกา ภาษีผล 2554; ศิริชัย กาญจนवासี, 2556; สุวิมล ติรกานันท์, 2551) ดังนั้น ในทางปฏิบัติ ค่าอำนาจจำแนกตั้งแต่ 0.20 ขึ้นไป ก็ถือว่าเพียงพอต่อการจำแนกกลุ่มสูงและกลุ่มต่ำได้ ในการวิจัยครั้งนี้ ผู้วิจัยนำค่าอำนาจจำแนกของข้อสอบแต่ละข้อมาหาเฉลี่ยเป็นค่าดัชนีมาตรฐานอำนาจจำแนก พบว่า มีจำนวน 161 ค่า มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนก (M) เท่ากับ .460 ส่วนเบี่ยงเบนมาตรฐาน (SD) เท่ากับ .137

จากผลการวิเคราะห์คุณลักษณะงานวิจัยที่มีผลต่อค่าดัชนีมาตรฐานคุณภาพเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ไม่พบตัวแปรคุณลักษณะใดที่มีอิทธิพลต่อทุกดัชนีมาตรฐานทั้งความเที่ยง ความตรง ความยาก และอำนาจจำแนก ซึ่งการอภิปรายเกี่ยวกับคุณลักษณะงานวิจัยที่มีอิทธิพลต่อดัชนีมาตรฐาน ผู้วิจัยนำเสนอการอภิปรายในตัวแปรคุณลักษณะงานวิจัยที่เกี่ยวข้องสัมพันธ์โดยตรงกับการสร้างเครื่องมือวัด ดังนี้

ตัวแปรประเภทของเครื่องมือวัด พบว่า เป็นตัวแปรคุณลักษณะงานวิจัยที่มีอิทธิพลต่อดัชนีมาตรฐานทั้งความเที่ยง ความยาก และอำนาจจำแนก โดยการเปรียบเทียบค่าเฉลี่ยรายคู่พบว่า หากเป็นแบบสังเกตจะมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่าแบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า รวมถึงหากเป็นแบบสอบอัตนัย มาตรฐานประมาณค่า และแบบสังเกต จะมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่าแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า ทั้งนี้จะเป็นเพราะว่า ในกรณีของแบบสังเกต การที่ค่าดัชนีมาตรฐานความเที่ยงมีค่าสูงนั้น อาจเป็นเพราะในการประมาณค่าความเที่ยงของแบบสังเกต ผู้วิจัยมักใช้วิธีการประมาณค่าความเที่ยงที่เรียกว่า ความสอดคล้องระหว่างผู้ประเมิน (inter-rater reliability หรือ inter-rater agreement) ซึ่งจำเป็นต้องมีการพัฒนาและใช้เกณฑ์การให้คะแนนแบบรูบริก (scoring rubric) ประกอบด้วย ดังนั้น เพื่อยืนยันความสอดคล้องของผลการประเมินระหว่างผู้ประเมินตั้งแต่ 2 คนขึ้นไป โดยคำนวณสัดส่วนของผู้รับการประเมินที่ผู้ประเมิน 2 คนให้คะแนนตรงกัน ต่อจำนวนผู้รับการประเมินทั้งหมด (กมลวรรณ ตังธนาภานนท์, 2557; Nitko & Brookhart, 2007; Reynolds, et al., 2009) ดังนั้น ความเที่ยงประเภทนี้จะมีค่าสูงก็ต่อเมื่อผู้ประเมินซึ่งเป็นผู้ใช้เครื่องมือมีการพูดคุยถึงพฤติกรรมที่ทำการประเมิน และทำความเข้าใจที่มีต่อเกณฑ์การประเมินที่ตรงกันจะทำให้ค่าความเที่ยงสูงขึ้นไป (Miller, et al, 2009) ในกรณีของแบบสอบอัตนัยที่พบว่ามีความเที่ยงสูงกว่าแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่านั้น ขัดแย้งกับ สุวิมล ติรกานันท์ (2551) ที่กล่าวถึงปัญหาของการใช้แบบสอบอัตนัยว่า การให้คะแนนมักมีความเที่ยงต่ำ เพราะมีปัจจัยที่มีอิทธิพลต่อคะแนนได้มาก เช่น ความเห็นส่วนตัว อารมณ์ของผู้ตรวจ ความลำเอียง การใช้ภาษาเขียน ลำดับข้อความ เป็นต้น ทั้งนี้ ความเที่ยงต่ำอาจจะเกิดขึ้นได้จากผู้สอบและผู้ตรวจ กล่าวคือ ในการสอบแต่ละครั้งด้วยข้อสอบเดิม ผู้สอบจะตอบแตกต่างกัน ในขณะที่เดียวกัน ผู้ตรวจแต่ละคนจะให้คะแนนแตกต่างกันไป ทำให้ผลการประมาณค่าความเที่ยงของเครื่องมือต่ำ อย่างไรก็ตาม เมื่อพิจารณาจากเกณฑ์ค่าความเที่ยงของเครื่องมือวัดต่าง ๆ บุญใจ ศรีสถิตนรากร (2555) กล่าว

ว่า แบบสังเกตควรมีค่าความเที่ยงมากกว่า .80 ในขณะที่ สุวิมล ติรภานันท์ (2551) กล่าวว่า สำหรับการวิจัยพื้นฐานทั่วไป แบบสอบควรมีค่าสัมประสิทธิ์ความเที่ยงไม่ต่ำกว่า 0.70 และสำหรับการวิจัยที่สนใจขนาดของสหสัมพันธ์หรือความแตกต่างของค่าเฉลี่ยระหว่างกลุ่ม ซึ่งเป็นเครื่องประเภมาตรประมาณค่าควรมีค่าสัมประสิทธิ์ความเที่ยงที่ระดับ 0.80 ในขณะที่ Miller, et al. (2009) กล่าวว่า สำหรับแบบสอบมาตรฐาน (standardized tests) ทั้งการวัดผลสัมฤทธิ์และทัศนคติ ควรมีค่าสัมประสิทธิ์ความเที่ยงหรือความคงที่ในช่วง 0.80 ดังนั้นจึงเห็นได้ว่าในกรณีของแบบสังเกต และมาตรประมาณค่าจะมีการกำหนดค่าความเที่ยงขั้นต่ำสูงกว่าเครื่องมือประเภทแบบสอบ สำหรับกรณีของค่าดัชนีมาตรฐานความยาก ผลการวิจัยพบว่าแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า มีค่าเฉลี่ยดัชนีความยากสูงกว่าแบบสอบหลายตัวเลือกแบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า แบบสอบอัตนัย แบบสอบภาคปฏิบัติ และแบบสอบอัตนัยประยุกต์ (MEQ) ทั้งนี้ อาจจะเป็นเพราะแบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า เป็นเครื่องมือวัดที่นิยมนำมาใช้โดยมีกำหนดให้มีสถานการณ์ที่คล้ายคลึง หรือเลียนแบบเหตุการณ์ต่าง ๆ ในชีวิตจริง เพื่อให้ผู้สอบได้สัมผัสกับสภาพการณ์เหล่านั้น แล้วแสดงพฤติกรรมตอบสนองที่สังเกตได้ หรือวัดได้ออกมา โดย พิชิต ฤทธิ์จรูญ (2556) กล่าวว่า คำถามที่ใช้ถามมี 2 ลักษณะ คือ 1) ถามให้ประเมินสถานการณ์ดังกล่าวเพื่อตัดสินว่าควร-ไม่ควร ดี-ไม่ดี ทำ-ไม่ทำ ถูกต้อง-ไม่ถูกต้อง ใช้ได้-ใช้ไม่ได้ และรวมถึงกรณีที่ไม่อาจตัดสินใจได้ด้วย และ 2) ถามให้ระบุแนวทางที่ตนเองจะปฏิบัติ ถ้าหากตนเองเป็นบุคคลในสถานการณ์นั้น หรือเกี่ยวข้องในสถานการณ์นั้นจะปฏิบัติอย่างไร ซึ่งจะมีการให้คะแนนตามความเหมาะสมที่ผู้สอบเลือกคำตอบ จึงเป็นการวัดกับตัวแปรด้านสมรรถภาพทางสมอง และด้านจิตพิสัย จึงมีค่าเฉลี่ยความยากสูงกว่าแบบสอบประเภทอื่น กล่าวคือ เป็นแบบสอบที่ค่อนข้างง่ายกว่าแบบสอบประเภทอื่น ๆ ในส่วนของค่าดัชนีมาตรฐานอำนาจจำแนก พบการวิจัยพบว่า แบบสอบอัตนัย และแบบมาตรประมาณค่า มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าแบบสอบภาคปฏิบัติ นอกจากนี้ แบบมาตรประมาณค่า มีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าแบบสอบอัตนัยประยุกต์ (MEQ) จะเห็นได้ว่าแบบมาตรประมาณค่ามีค่าอำนาจจำแนกสูงกว่าเครื่องมือวัดอีก 2 ประเภท ทั้งนี้อาจเป็นเพราะลักษณะการตอบของแบบมาตรประมาณค่า ที่เป็นการตอบตามการรับรู้ของผู้ตอบ ในขณะที่แบบสอบภาคปฏิบัติ และแบบสอบอัตนัยประยุกต์ (MEQ) เป็นเครื่องมือที่กำหนดให้ผู้สอบหาคำตอบ หรือแสดงทักษะปฏิบัติภายใต้เงื่อนไขหรือสถานการณ์ที่กำหนด นอกจากนี้ สูตรที่ใช้ในการประมาณค่าอำนาจจำแนกนั้นมีความแตกต่างกัน กล่าวคือ สำหรับแบบมาตรประมาณค่าจากงานวิจัยที่นำมาสังเคราะห์เกือบทั้งหมด (ร้อยละ 91.30) ใช้ค่า Item-total correlation ในการประมาณค่าอำนาจจำแนก ในขณะที่แบบสอบภาคปฏิบัติ และแบบสอบอัตนัยประยุกต์ (MEQ) ใช้สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำในการประมาณค่าอำนาจจำแนก ทั้งนี้ มีนักวิชาการบางส่วนได้แสดงความเห็นว่าการใช้ค่า Item-total correlation ในการหาค่าอำนาจจำแนกของแบบมาตรประมาณค่ายังขาดความเหมาะสม ดังเช่น บุญเรียง ขจรศิลป์ (2550) ได้กล่าวว่า ในกรณีที่คำตอบแต่ละข้อ มีข้อมูลเป็น 0 หรือ 1 ผลการวิเคราะห์ที่ได้จาก Item-total correlation จะให้ผลการวิเคราะห์เท่ากับค่าสหสัมพันธ์แบบ Point-biserial (r_{pbis}) ซึ่งคือ ค่าอำนาจจำแนก แต่หากข้อมูลที่ไม่ใช่ 0 หรือ 1 เช่น แบบมาตรประมาณค่า ค่า item-total correlation คือ ค่าความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมทั้งฉบับ เป็นสิ่งที่สะท้อนถึงความสอดคล้องหรือความคงที่ของข้อคำถามแต่ละข้อกับคะแนนรวมทั้งฉบับ ไม่ใช่ค่าอำนาจจำแนก ดังนั้น ตัวประมาณค่าที่ต่างกันอาจจะเป็นสาเหตุที่ทำให้ค่าดัชนีมาตรฐานอำนาจจำแนกจากแบบสอบต่างประเภทกันมีค่าแตกต่างกัน

ตัวแปรประเภทจำนวนตัวเลือก/เกณฑ์ พบว่า เป็นตัวแปรคุณลักษณะงานวิจัยที่มีอิทธิพลต่อค่าดัชนีมาตรฐานความเที่ยง และความยาก โดยการเปรียบเทียบค่าเฉลี่ยรายคู่ พบว่า เครื่องมือวัดที่มี

จำนวนมากกว่า 5 ตัวเลือก/เกณฑ์ จะมีค่าดัชนีมาตรฐานความเที่ยงมากกว่าเครื่องมือวัดที่มีจำนวน 3 หรือ 4 ตัวเลือก/เกณฑ์ นอกจากนี้ยังพบอีกว่า เครื่องมือวัดที่ไม่มีตัวเลือก/เกณฑ์ จะมีค่าดัชนีมาตรฐานความเที่ยงมากกว่าเครื่องมือวัดที่มีจำนวน 4 ตัวเลือก/เกณฑ์ ซึ่งผลการวิจัยไม่สอดคล้องกับผลวิจัยของ ไชยยศ ไพวิทยศิริธรรม (2544) ได้ทำการวิจัยเรื่อง การวิเคราะห์ห่อภิมาณคุณภาพของแบบสอบหลายตัวเลือกที่มีรูปแบบของแบบสอบ บริบทของแบบสอบ และวิธีการตรวจให้คะแนนต่างกัน ผลการวิจัยพบว่า ตัวแปรจำนวนตัวเลือกไม่มีอิทธิพลต่อค่าดัชนีมาตรฐานความเที่ยง ในขณะที่งานวิจัยของ อาภาพรรณ ประทุมไทย (2553) ที่ได้วิเคราะห์ห่อภิมาณงานวิจัยที่เกี่ยวกับการสร้างแบบวัดความถนัด ที่พบว่า เครื่องมือวัดที่ไม่ระบุจำนวนตัวเลือกมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่าเครื่องมือวัดที่มีจำนวน 4 หรือ 5 ตัวเลือก ในส่วนของค่าดัชนีมาตรฐานความยากผลการวิจัยพบว่า เครื่องมือที่มีตัวเลือกมากกว่า 1 ประเภท จะมีค่าเฉลี่ยดัชนีมาตรฐานความยากสูงกว่าเครื่องมือวัดที่ไม่มีตัวเลือก/เกณฑ์ และเครื่องมือวัดที่มี 4 ตัวเลือก/เกณฑ์ ผลการวิจัยไม่สอดคล้องกับผลวิจัยของ ไชยยศ ไพวิทยศิริธรรม (2544) และงานวิจัยของ อาภาพรรณ ประทุมไทย (2553) ที่พบว่า ตัวแปรจำนวนตัวเลือกไม่มีอิทธิพลต่อค่าดัชนีมาตรฐานความยาก

ตัวแปรการตรวจให้คะแนน พบว่า เป็นตัวแปรคุณลักษณะงานวิจัยที่มีอิทธิพลต่อค่าดัชนีมาตรฐานความเที่ยง โดยเครื่องมือวัดที่มีการให้คะแนนมากกว่าสองค่า จะมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่าเครื่องมือวัดที่มีการให้คะแนนแบบสองค่า นอกจากนี้ยังพบว่า เครื่องมือวัดที่มีการให้คะแนนแบบสองค่าและมากกว่าสองค่า จะมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่าเครื่องมือวัดที่มีการให้คะแนนแบบสองค่า หรือให้คะแนนมากกว่าสองค่า ทั้งนี้อาจเนื่องมาจากความแปรปรวนของข้อมูล (data variation) ดังที่ บุญใจ ศรีสถิตนรากร (2555) กล่าวว่าข้อมูลที่มีการกระจายแบบปกติ ค่าความเที่ยงจะสูงกว่าข้อมูลที่มีการกระจายแบบเบ้ รวมถึงข้อมูลที่ได้จากการมีความแปรปรวนมากจะทำให้ค่าความเที่ยงของเครื่องมือวัดมีค่าเข้าใกล้ 1

ตัวแปรจำนวนข้อคำถามในแต่ละฉบับ พบว่า เป็นตัวแปรคุณลักษณะงานวิจัยที่มีอิทธิพลต่อค่าดัชนีมาตรฐานความเที่ยง และอำนาจจำแนก โดย เครื่องมือวัดที่มีจำนวนข้อคำถามในแต่ละฉบับมากกว่า 30 ข้อ ขึ้นไป จะมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่าเครื่องมือวัดที่มีจำนวนข้อคำถามในแต่ละฉบับน้อยกว่าหรือเท่ากับ 30 ข้อ ทั้งนี้ เป็นเพราะความยาวของแบบสอบเป็นปัจจัยสำคัญที่ส่งผลต่อค่าความเที่ยง เนื่องมาจากความยาวของแบบสอบมีผลต่อทั้งความแปรปรวนของคะแนนจริง และความแปรปรวนของคะแนนที่สังเกตได้ การเพิ่มความยาวของแบบสอบจะช่วยเพิ่มความแปรปรวนของคะแนนจริงในอัตราที่รวดเร็วกว่าการเพิ่มความแปรปรวนของคะแนนความคลาดเคลื่อน จึงทำให้ค่าสัมประสิทธิ์ความเที่ยงสูงขึ้น (ศิริชัย กาญจนวาสี, 2556; Reynolds, et al., 2009) อย่างไรก็ตาม ความยาวของแบบสอบหรือเครื่องมือที่พัฒนาขึ้นนั้นมีผลต่อค่าสัมประสิทธิ์ความเที่ยง แต่การเพิ่มข้อคำถามเพื่อให้ค่าสัมประสิทธิ์ความเที่ยงสูงขึ้นนั้นไม่ใช่สิ่งที่พึงกระทำ เนื่องจากจำนวนข้อคำถามในแบบสอบหรือเครื่องมือที่พัฒนาขึ้นนั้นได้ถูกกำหนดตามโครงสร้างที่ออกแบบไว้แล้ว และเป็นผลสืบเนื่องมาจากความตรงตามเนื้อหาและความตรงตามโครงสร้าง ซึ่งเป็นคุณภาพเครื่องมือที่สำคัญ การเพิ่มข้อสอบหรือข้อคำถามอย่างไม่มีเหตุผลอาจส่งผลต่อความตรงทั้งสองได้ ประกอบกับการที่หรือเครื่องมือที่พัฒนาขึ้นนั้นมีสัมประสิทธิ์ความเที่ยงสูง ไม่จำเป็นต้องเป็นเครื่องมือที่มีคุณภาพเสมอไป แต่อาจเกิดจากจำนวนข้อคำถามที่มากเกินไป หรือคำถามที่มีความหมายซ้ำซ้อน (redundancy) (บุญใจ ศรีสถิตนรากร, 2555; ศิริชัย กาญจนวาสี, 2556; สุวิมล ติรกันันท์, 2551) ทั้งนี้ ผลการวิจัยดังกล่าวสอดคล้องกับผลการวิจัยของ ไชยยศ ไพวิทยศิริธรรม (2544) และงานวิจัยของ อาภาพรรณ ประทุมไทย (2553) ที่พบว่า เครื่องมือวัดที่มีจำนวนข้อคำถามจำนวนข้อมากกว่า จะมีค่าเฉลี่ยดัชนีมาตรฐานความเที่ยงสูงกว่าเครื่องมือวัดที่มีจำนวนข้อคำถามน้อยกว่า อย่างไรก็ตามเมื่อ

พิจารณาในมิติของค่าดัชนีมาตรฐานอำนาจจำแนก กลับพบว่า เครื่องมือวัดที่มีจำนวนข้อคำถามในแต่ละฉบับน้อยกว่าหรือเท่ากับ 30 ข้อ จะมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าเครื่องมือวัดที่มีจำนวนข้อคำถามในแต่ละฉบับมากกว่า 30 ข้อ ขึ้นไป ผลการวิจัยดังกล่าวสอดคล้องงานวิจัยของ อาภาพรธน์ ประทุมไทย (2553) ที่พบว่า เครื่องมือวัดที่มีจำนวนข้อคำถามน้อยกว่าจะมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าเครื่องมือวัดที่มีจำนวนข้อคำถามจำนวนข้อมากกว่า ทั้งนี้อาจจะเป็นเพราะจำนวนข้อคำถามที่มากเกินไปอาจส่งผลต่อความตั้งใจของผู้สอบ หรือผู้ตอบเครื่องมือวัดนั้น ๆ ทำให้ค่าอำนาจจำแนกลดต่ำลง

ตัวแปรจำนวนผู้สอบในแต่ละฉบับ พบว่า เป็นตัวแปรคุณลักษณะงานวิจัยที่มีอิทธิพลต่อค่าดัชนีมาตรฐานอำนาจจำแนก โดยเครื่องมือวัดที่มีจำนวนผู้สอบในแต่ละฉบับมากกว่า 200 คน ขึ้นไป จะมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าเครื่องมือวัดที่มีจำนวนผู้สอบในแต่ละฉบับน้อยกว่าหรือเท่ากับ 200 คน ผลการวิจัยดังกล่าวสอดคล้องกับผลการวิจัยของ ไชยยศ ไพวิทยศิริธรรม (2544) และงานวิจัยของ อาภาพรธน์ ประทุมไทย (2553) ที่พบว่า เครื่องมือวัดที่มีจำนวนผู้สอบในแต่ละฉบับมากกว่า 200 คน ขึ้นไป จะมีค่าเฉลี่ยดัชนีมาตรฐานอำนาจจำแนกสูงกว่าเครื่องมือวัดที่มีจำนวนผู้สอบในแต่ละฉบับน้อยกว่าหรือเท่ากับ 200 คน อย่างไรก็ตาม ศิริชัย กาญจนวาสิ (2556) ได้เสนอแนะว่าการนำเครื่องมือวัดไปทดลองใช้ ควรมีตัวอย่างไม่น้อยกว่า 50 คน หากเป็นไปได้ควรมีขนาดเป็นร้อย เพื่อให้ผลการวิเคราะห์ค่อนข้างคงที่และน่าเชื่อถือ รวมถึงควรมีการคัดเลือกตัวอย่างเป็นไปด้วยความระมัดระวังที่มีความเป็นตัวแทนของกลุ่มที่ต้องการศึกษา

3. องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

3.1 จากการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัด ซึ่งเป็นการศึกษาในเชิงคุณภาพ พบว่า เครื่องมือวัดความรู้เรื่องพลเมืองมีงานวิจัยที่ศึกษาเพียง 1 เรื่อง ซึ่งทำให้ไม่เห็นสารสนเทศที่มากพอ ทั้งนี้ สาเหตุที่ทำให้มีผู้สนใจศึกษาเรื่องนี้จำนวนน้อยอาจเนื่องมาจาก ความรู้เรื่องพลเมืองมีความคล้ายคลึงกับเนื้อหาสาระในกลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 จึงไม่จำเป็นต้องสร้างเครื่องมือเฉพาะขึ้นมาวัด นอกจากนี้ สมรรถนะเกี่ยวกับจิตสำนึกต่อโลก พบว่า มีผู้ศึกษาจำนวนน้อยเช่นกัน

3.2 จากการพิจารณาชื่อแบบวัดที่นำมาสังเคราะห์ พบว่า แบบวัดทักษะชีวิต เป็นแบบวัดที่มีสมรรถนะที่สอดคล้องกับสมรรถนะในศตวรรษที่ 21 มากที่สุด ครอบคลุมหลายสมรรถนะ ซึ่งองค์การอนามัยโลก (World Health Organization, 1997) ให้นิยามทักษะชีวิตว่า เป็นความสามารถในการปรับตัว และมีพฤติกรรมไปในทางที่ถูกต้อง ในการเผชิญกับสิ่งท้าทายต่าง ๆ ที่เกิดขึ้นในชีวิตประจำวันได้อย่างมีประสิทธิภาพ ซึ่งจำแนกออกเป็น 10 องค์ประกอบ ได้แก่ ทักษะการตัดสินใจ ทักษะการแก้ปัญหา การมีความคิดริเริ่มสร้างสรรค์ การมีความคิดวิจารณ์ญาณ ทักษะการสื่อสารได้อย่างมีประสิทธิภาพ ทักษะการสร้างสัมพันธ์ภาพระหว่างบุคคล ความสามารถที่จะตระหนักรู้ในตน ความเห็นอกเห็นใจผู้อื่น การจัดการเกี่ยวกับภาวะอารมณ์ และการจัดการกับความเครียด ซึ่งจะเห็นได้ว่าองค์ประกอบเหล่านี้ล้วนแล้วแต่มีความจำเป็นในการดำเนินชีวิตในศตวรรษที่ 21 และสอดคล้องกับองค์ประกอบทั้ง 11 องค์ประกอบที่ผู้วิจัยผู้วิจัยสังเคราะห์ขึ้น ซึ่งประกอบไปด้วย จิตสำนึกต่อโลก ความรู้เรื่องพลเมือง การคิดอย่างมีวิจารณ์ญาณ การแก้ปัญหา การสร้างสรรค์และนวัตกรรม การสื่อสาร ทักษะด้านการทำงานเป็นทีม ความร่วมมือ ภาวะผู้นำ ความรับผิดชอบ การรู้เท่าทันสื่อและสารสนเทศ และการรู้เท่าทันเทคโนโลยี

3.3 จากการสังเคราะห์องค์ความรู้ พบว่า ประเภทเครื่องมือที่นิยมนำมาใช้วัดสมรรถนะในศตวรรษที่ 21 ได้แก่ แบบวัดเชิงสถานการณ์ และ มาตราประมาณค่า เนื่องจากแบบวัดเชิงสถานการณ์

สามารถวัดความรู้ขั้นสูงทั้งด้านสมรรถภาพทางสมองและด้านจิตพิสัย (พิชิต ฤทธิ์จรูญ, 2556) ซึ่งสอดคล้องกับลักษณะของสมรรถนะในศตวรรษที่ 21 ที่มีทั้งทักษะการคิดขั้นสูง และตัวแปรด้านจิตพิสัย นอกจากนั้นแบบวัดเชิงสถานการณ์มีจุดเด่น คือ การเข้าใจผู้ตอบให้ติดตาม เพราะได้อ่านเรื่องราวและได้คิดมากกว่าข้อสอบประเภทอื่น ๆ สำหรับมาตรฐานค่า มีจุดเด่นที่อำนวยความสะดวกให้ผู้ให้หลายประการ เช่น 1) สร้างง่าย ประหยัด สะดวกในการนำไปใช้ การให้คะแนนและการวิเคราะห์ ผล 2) ไม่ต้องใช้กลุ่มผู้เชี่ยวชาญในการตัดสินค่าคะแนนประจำข้อ 3) ไม่ต้องคำนวณค่าประจำข้อ 4) มีความเที่ยงสูงมาก ใช้ข้อคำถามเพียงไม่กี่ข้อก็มีความเที่ยงสูงพอ ๆ กับวิธีการอื่น ๆ ที่ใช้จำนวนข้อมาก 5) สามารถนำมาประยุกต์ใช้ในการวัดเจตคติที่มีต่อสิ่งต่าง ๆ สถานการณ์ต่าง ๆ ได้อย่างกว้างขวาง ทั้งยังสามารถวัดทิศทาง และปริมาณความมากน้อยของเจตคติได้อีกด้วย (Kubiszyn and Borichm, 1984; วิเชียร เกตุสิงห์, 2530; บุญธรรม กิจปรีดา บริสุทธิ, 2531; ศักดิ์ สุนทรเสณี, 2531 อ้างถึงใน ภัคณัฐ สมพงษ์ธรรม, 2551) นอกจากนั้น จากงานวิจัยของ ภัคณัฐ สมพงษ์ธรรม (2551) ที่เปรียบเทียบคุณภาพของแบบวัดความสามารถในการเผชิญและฟื้นฝ่าอุปสรรคตามทฤษฎีของสโตสซ์ระหว่างมาตรฐานค่ากับแบบวัดชนิดสถานการณ์: การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค พบว่า เมื่อเปรียบเทียบโมเดลการวัดชนิดมาตรฐานค่าและโมเดลการวัดของแบบวัดชนิดสถานการณ์มีความสอดคล้องกับข้อมูลเชิงประจักษ์เหมือนกัน และมีความตรงตามสภาพไม่แตกต่างกัน ทั้งนี้ เมื่อพิจารณาความเที่ยงพบว่า แบบวัดชนิดสถานการณ์มีความเที่ยงสูงกว่ามาตรฐานค่า ในช่วงระดับความสามารถในการเผชิญและฟื้นฝ่าอุปสรรคต่ำถึงค่อนข้างสูง ($-4 \leq 0 \leq 2$) แบบวัดชนิดสถานการณ์ให้ฟังก์ชันสารสนเทศสูงกว่าแบบวัดชนิดมาตรฐานค่า ส่วนในช่วงระดับความสามารถในการเผชิญและฟื้นฝ่าอุปสรรคสูง ($3 \leq 0 \leq 4$) แบบวัดชนิดมาตรฐานค่าให้ฟังก์ชันสารสนเทศสูงกว่าแบบวัดชนิดสถานการณ์ เมื่อพิจารณาประสิทธิภาพสัมพัทธ์เฉลี่ยพบว่าแบบวัดชนิดสถานการณ์มีประสิทธิภาพสูงกว่าแบบวัดชนิดมาตรฐานค่า

สำหรับเครื่องมือวัดสมรรถนะด้านการสร้างสรรค์และนวัตกรรม มีความแตกต่างจากสมรรถนะอื่น ๆ เนื่องจาก การวัดสมรรถนะด้านการสร้างสรรค์และนวัตกรรม ไม่สามารถใช้เครื่องมือที่มีลักษณะเป็นตัวเลือกให้เลือกตอบ เพราะถือเป็นการปิดกั้นความคิดสร้างสรรค์ งานวิจัยส่วนใหญ่ (สุชาติ ยอดสุรงค์, 2552 ; ศศิธร เวียงอินทร์, 2547; กชพรรณ เกษัชชา, 2560) จึงใช้วิธีการเขียนตอบ และสร้างเกณฑ์การให้คะแนน โดยพิจารณาตามแนวคิดของ Torrance ซึ่งประกอบด้วย ความคิดคล่อง ความคิดยืดหยุ่น ความคิดริเริ่ม และความคิดละเอียดลออ (สมศักดิ์ ภูวิภาดารัตน์, 2544)

ข้อเสนอแนะในการวิจัย

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ผลการวิจัยครั้งนี้ทำให้ทราบว่าในภาพรวมของงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา มีคุณภาพอยู่ในระดับดี แต่มีบางประเด็นที่ควรให้ความสำคัญมากขึ้นโดยเฉพาะประเด็นที่มีคุณภาพอยู่ในระดับค่อนข้างต่ำและระดับปานกลาง เพื่อเพิ่มคุณค่าของงานวิจัยดังนี้

1.1 สำหรับนักวิจัย หรือนิสิต นักศึกษาผู้ซึ่งอยู่ระหว่างพัฒนาหัวข้อวิจัย ควรให้ความสำคัญกับการทบทวนเอกสารและงานวิจัยที่เกี่ยวข้อง ถึงแม้ในภาพรวมจะมีผลการประเมินคุณภาพอยู่ในระดับดี แต่เมื่อพิจารณาจากค่าเฉลี่ยแล้ว เป็นด้านที่มีค่าเฉลี่ยต่ำที่สุด ประกอบผลการประเมินยังพบว่า ความทันสมัยของเอกสารและงานวิจัยที่เกี่ยวข้องอยู่ในระดับค่อนข้างต่ำ รวมถึงเอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย และสัดส่วนของเอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทย

และภาษาต่างประเทศ ยังอยู่ในระดับปานกลาง การทบทวนเอกสารและงานวิจัยที่เกี่ยวข้องเป็นจุดเริ่มต้นของการทำวิจัยที่ได้มุมมองที่แปลกใหม่ ลดความซ้ำซ้อนกับงานวิจัยที่ผ่านมา

1.2 ในการดำเนินการวิจัยเพื่อพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ผู้วิจัยควรระบุเหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน สามารถนำผลที่ได้จากการวิจัยใช้ประโยชน์ได้ในทางวิชาการ เนื่องจากในการพัฒนาผู้เรียนให้มีสมรรถนะในศตวรรษที่ 21 นั้น เครื่องมือวัดถือเป็นสิ่งสำคัญในการวัดเพื่อให้ได้สารสนเทศ หรือข้อมูลพื้นฐานสำหรับการวางแผนพัฒนาส่งเสริมให้ผู้เรียนมีสมรรถนะในศตวรรษที่ 21 หรือใช้สำหรับวัดผลสำเร็จที่เกิดขึ้นการทำกิจกรรมส่งเสริมสมรรถนะดังกล่าว ดังนั้น การกำหนดจุดมุ่งหมาย และการนำเครื่องมือวัดที่ได้จากการวิจัยไปใช้ในเชิงวิชาการจึงเป็นเรื่องสำคัญที่ไม่ควรละเลย

2. สำหรับหน่วยงานที่ต้องการคัดเลือกเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ที่มีอยู่แล้วไปใช้หรือต้องการพัฒนาเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ฉบับใหม่ ควรศึกษารายละเอียดเครื่องมือที่เหมาะสมในแต่ละสมรรถนะ เนื่องจาก มีความแตกต่างกัน ทั้งในด้านประเภทเครื่องมือ และการตรวจให้คะแนน

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. เพื่อให้การดำเนินการวิจัยเกี่ยวกับการพัฒนาเครื่องมือวัดสมรรถนะให้มีความสมบูรณ์มากขึ้น ผู้วิจัยขอให้ข้อเสนอแนะดังนี้

1.1 นักวิจัยที่พัฒนาเครื่องมือวัดควรรายงานค่าดัชนีคุณภาพเครื่องมือทั้งค่าความยาก และอำนาจจำแนกเป็นรายข้อให้ครบถ้วน เพื่อจะเป็นประโยชน์ต่อนักวิจัยอื่น ๆ ในการนำข้อมูลมาใช้สังเคราะห์ต่อไป

1.2 ในการตรวจสอบคุณภาพเครื่องมือด้วยดัชนีคุณภาพต่าง ๆ ควรพิจารณาความเหมาะสมของวิธีการที่เลือกมา เช่น จากการศึกษาพบว่ามีงานวิจัยบางส่วนเลือกใช้การเปรียบเทียบคะแนนกับกลุ่มที่ทราบผล (known group) สำหรับการตรวจสอบความตรงเชิงโครงสร้างของเครื่องมือวัด โดยนักวิจัยได้สร้างเครื่องมืออีก 1 ฉบับ เพื่อให้ครูเป็นผู้ประเมินนักเรียนสำหรับการจำแนกกลุ่ม หรือการใช้คะแนนเฉลี่ยสะสมในการจำแนกกลุ่ม ซึ่งไม่สอดคล้องกับสมรรถนะที่ต้องการพัฒนาเครื่องมือวัด ทำให้การประมาณค่าความตรงดังกล่าวมีความคลาดเคลื่อน

1.3 จากการศึกษาพบว่า การตรวจสอบคุณภาพเครื่องมือวัดในด้านความตรง งานวิจัยทุกเรื่องมีการตรวจสอบความตรงเชิงเนื้อหา โดยการตัดสินจากผู้เชี่ยวชาญ และการตรวจสอบความตรงเชิงโครงสร้างด้วยการวิเคราะห์องค์ประกอบเชิงยืนยันมากที่สุด อย่างไรก็ตาม ในการตรวจสอบความตรงซึ่งมีแนวคิดเอกลักษณ์ (unitary concept) กล่าวคือ ความตรงเป็นคุณลักษณะเดียวที่จำเป็นต้องตรวจสอบด้วยวิธีการหลายวิธี เพื่อให้ได้หลักฐานต่าง ๆ สำหรับสนับสนุนหรือยืนยันความตรง ดังนั้นควรมีการตรวจสอบความตรงในมิติของความตรงตามเกณฑ์สัมพันธ์ (criterion-related validity) เพิ่มเติมจะทำให้การตรวจสอบความตรงมีความสมบูรณ์ยิ่งขึ้น

2. เนื่องจากการวิจัยในครั้งนี้ดำเนินการภายใต้ขอบเขตการวิจัยว่าต้องเป็นงานวิจัยที่ใช้ทฤษฎีการทดสอบแบบดั้งเดิม เพื่อนำดัชนีมาตรฐานคุณภาพเครื่องมือมาใช้สำหรับการวิเคราะห์ถ้อยความ อย่างไรก็ตาม เนื่องจากในอนาคตมีแนวโน้มของงานวิจัยที่ใช้ทฤษฎีการทดสอบแนวใหม่เพิ่มมากขึ้น จึงควรศึกษาวิธีการสังเคราะห์คุณภาพเครื่องมือวัดที่ใช้กับทฤษฎีการทดสอบแนวใหม่ หรือกรณีที่ใช้หลายทฤษฎีร่วมกัน

3. จากผลการสังเคราะห์พบว่า งานวิจัยที่พัฒนาเครื่องมือวัดไม่ได้เป็นเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 โดยตรง แต่สมรรถนะหลัก 4 ด้าน โดยประกอบด้วยสมรรถนะย่อย 11 ด้านที่ผู้วิจัยใช้เป็นกรอบในการสังเคราะห์ครั้งนี้ ไปปรากฏเป็นองค์ประกอบย่อยของเครื่องมือวัดอื่น ๆ เช่น แบบวัดคุณลักษณะอันพึงประสงค์ แบบวัดทักษะชีวิต แบบวัดความคิดสร้างสรรค์ เป็นต้น ทำให้ตัวบ่งชี้หรือพฤติกรรมที่ทำการวัดยังไม่สะท้อนแนวคิดของสมรรถนะในศตวรรษที่ 21 เท่าที่ควร เช่น สมรรถนะการสร้างสรรคและนวัตกรรม (creativity and innovation) ยังไม่ปรากฏเป็นที่ชัดเจนมากนัก เนื่องจากแบบวัดที่ผ่านมาเน้นเพื่อการวัดความคิดสร้างสรรค์โดยเฉพาะ หรือในสมรรถนะจิตสำนึกต่อโลก (global awareness) เป็นการวัดผ่านเรื่องอื่น เช่น จิตสำนึกต่อภาวะโลกร้อน เป็นต้น หรือความรู้เรื่องพลเมือง (civic literacy) ที่ยังคงมีงานวิจัยที่พัฒนาเครื่องมือวัดในสมรรถนะนี้เพียงเรื่องเดียว ดังนั้น ในการวิจัยเพื่อพัฒนาเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ระดับมัธยมศึกษา ควรยึดกรอบสมรรถนะในศตวรรษที่ 21 เป็นกรอบแนวคิดหลักในการพัฒนา

รายการอ้างอิง

- กขพรรณ เกษัชชา. (2560). *การสร้างแบบวัดความคิดสร้างสรรค์ทางวิทยาศาสตร์สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3*. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
- กมลวรรณ ตังธนากานนท์. (2557). *การวัดและประเมินทักษะการปฏิบัติ*. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ขวัญยืน มูลศรี. (2548). *การสร้างแบบวัดทักษะชีวิตสำหรับนักเรียนช่วงชั้นที่ 3 (ชั้นมัธยมศึกษาปีที่ 1-3)*. (วิทยานิพนธ์มหาบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, บัณฑิตวิทยาลัย, สาขาวิชาการวัดผลการศึกษา.
- คณะกรรมการสภာวิจัยแห่งชาติ สาขาการศึกษา. (2555). *หลักสูตร “มโนทัศน์ที่คาดเคลื่อนในการวิจัยการศึกษา” เอกสารในโครงการ “การพัฒนาหลักสูตรการฝึกอบรมนักวิจัยการศึกษาเพื่อเพิ่มศักยภาพด้านการวิจัย”*. กรุงเทพมหานคร: ห้างหุ้นส่วนจำกัดไอออนพรีนติ้ง.
- โชติกา ภาชีผล. (2554). *การสร้างและพัฒนาเครื่องมือในการวัดและประเมินผลการศึกษา*. พิมพ์ครั้งที่ 2. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- โชติกา ภาชีผล, ณีฐกรรณ์ หลาวทอง และกมลวรรณ ตังธนากานนท์. (2558). *การวัดและประเมินผลการเรียนรู้*. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ไชยยศ ไพวิทยศิริธรรม. (2544). *การวิเคราะห์ห่อภิมาณคุณภาพของแบบสอบหลายตัวเลือกที่มีรูปแบบของแบบสอบ บริบทของแบบสอบ และวิธีการตรวจให้คะแนนต่างกัน*. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
- ณีฐกรรณ์ หลาวทอง. (2551). *การวัดและประเมินผลการศึกษา*. กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- ณีฐกรรณ์ หลาวทอง. (2559). *การสร้างเครื่องมือการวิจัยทางการศึกษา*. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- ทิพย์ฉมพร เกษโกมล. (2554). *การสร้างแบบวัดจิตลักษณะและพฤติกรรมแบบมาตราประมาณค่า*. วารสารพยาบาลตำรวจ, 3(1): 1-11
- ทัศน์ศิริรินทร์ สว่างบุญ. (2548). *การสังเคราะห์งานวิจัยด้านการมีส่วนร่วมของโรงเรียนกับชุมชน: การวิเคราะห์ห่อภิมาณและการวิเคราะห์กระบวนการทางปัญญาอภีมาณ*. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาวิจัยการศึกษา.
- นงลักษณ์ วิรัชชัย. (2542). *การวิเคราะห์ห่อภิมาณ (meta-analysis)*. กรุงเทพมหานคร: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- นิติบดี สุขเจริญ และ วัลย์พจน์ อยู่ในศีล. (2557). *การวิเคราะห์ห่อภิมาณ และการสังเคราะห์ห่อภิมาณ*. ว.มร.ม. (มนุษยศาสตร์และสังคมศาสตร์). 8(3): 43-56.
- น้ำทิพย์ องอาจวานิชย์. (2556). *การพัฒนาแบบวัดทักษะแห่งศตวรรษที่ ๒๑ ตามการรับรู้ของผู้เรียนมัธยมศึกษาตอนต้น: การประยุกต์ใช้แนวคิดการเข้าถึงคุณลักษณะที่มุ่งวัดของแบบสอบ*. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.

- บุญใจ ศรีสถิตยน์รากูร. (2555). *การพัฒนาและตรวจสอบคุณภาพเครื่องมือวิจัย : คุณสมบัติการวัดเชิงจิตวิทยา (Development and validation of research instruments : Psychometric properties)*. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย
- บุญธรรม กิจปรีดาบริสุทธิ์. (2553). *เทคนิคการสร้างเครื่องมือเก็บรวบรวมข้อมูลสำหรับการวิจัย*. พิมพ์ครั้งที่ 7. กรุงเทพฯ : คณะสังคมศาสตร์และมนุษยศาสตร์ มหาวิทยาลัยมหิดล.
- บุญเรียง ขจรศิลป์. (2550). *ทำวิจัยอย่างไรให้มีคุณภาพ: ความคลาดเคลื่อนในการทำงานวิจัยทางการศึกษา. ในการประชุมสัมมนาทางวิชาการกระบวนการวิจัยกับการพัฒนาทางการศึกษา*. กรุงเทพฯ: คณะศึกษาศาสตร์มหาวิทยาลัยเกษตรศาสตร์.
- ปวีณา มะแซ. (2561). *การพัฒนาแบบวัดทักษะการรู้เท่าทันสื่อในศตวรรษที่ 21 ของนักเรียนชั้นมัธยมศึกษาตอนต้นโดยประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค*. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยสงขลานครินทร์, คณะศึกษาศาสตร์, สาขาวิชาการวิจัยและประเมินผลการศึกษา.
- ปาริชาติ อังกาบ. (2561). *การสร้างแบบวัดทักษะชีวิตและอาชีพตามแนวคิดทักษะในศตวรรษที่ 21 สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย*. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
- พิชญา ตีมี. (2559). *การพัฒนาแนวทางการประเมินทักษะการเรียนรู้และนวัตกรรมของผู้เรียน ตามแนวคิดการประเมินผู้เรียนในศตวรรษที่ 21*. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). มหาวิทยาลัยนเรศวร, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา
- พิชิต ฤทธิ์จรรยา. (2556). *หลักการวัดและประเมินผลการศึกษา*. พิมพ์ครั้งที่ 8. กรุงเทพฯ: บริษัท แฮาส์ ออฟเครอมีส์ จำกัด.
- ภักดี ภูมิธรรม. (2551). *การเปรียบเทียบคุณภาพของแบบวัดความสามารถในการเผชิญและฟื้นตัวต่อความเครียดของบุคลากรระหว่างมาตรฐานค่ากับแบบวัดชนิดสถานการณ์ : การประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค*. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
- เยาวดี วิบูลย์ศรี. (2556). *การวัดและการสร้างแบบสอบผลสัมฤทธิ์*. พิมพ์ครั้งที่ 11. กรุงเทพฯ: สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- รัชนิกุล ภิญโญภานูวัฒน์. (2536). *การวิเคราะห์ห่อภิมานเพื่อสังเคราะห์งานวิจัยเกี่ยวกับการสร้าง และพัฒนาแบบทดสอบและแบบวัด*. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). คณะศึกษาศาสตร์, มหาวิทยาลัยเกษตรศาสตร์, สาขาวิชาวิจัยและประเมินผลการศึกษา.
- รัตนะ บัวสนธ์ และคณะ. (2557). *การประเมินคุณภาพวิทยานิพนธ์ของนิสิตระดับปริญญาโท บัณฑิตวิทยาลัยนเรศวร*. *วารสารศึกษาศาสตร์ มหาวิทยาลัยนเรศวร* 16(1): 120-126.
- วรรณดี แสงประทีปทอง. (2545). *การพัฒนาเครื่องมือสำหรับการประเมินทางการศึกษา*. นนทบุรี: โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช.
- วรรณดี แกมเกตุ. (2555). *วิธีวิทยาการวิจัยทางพฤติกรรมศาสตร์*. พิมพ์ครั้งที่ 3. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- วัยญา ยิ้มยวน (2547). *การวิเคราะห์ห่อภิมานของปัจจัยที่มีความสัมพันธ์กับการคิดวิจารณ์ญาณ*. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาวิจัย การศึกษา.

- ศศิธร เวียงอินทร์. (2547). *การพัฒนาแบบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนในสังกัด สำนักงานเขตพื้นที่การศึกษาร้อยเอ็ด เขต 1.* (วิทยานิพนธ์ปริญญาโทบริหารศึกษิต). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
- ศิริชัย กาญจนวาสี. (2556). *ทฤษฎีการทดสอบแบบดั้งเดิม.* พิมพ์ครั้งที่ 7. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สมนึก ภัททิยธนี. (2560). *การวัดผลการศึกษา.* พิมพ์ครั้งที่ 11. กทม: ประสานการพิมพ์.
- สมบุญ ชิตพงศ์. (ม.ป.ป.). *การวัดคุณลักษณะทางด้านจิตวิทยา.* กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ
- สมศักดิ์ ภู่วิภาดาธารธณ์. (2544). *เทคนิคการส่งเสริมความคิดสร้างสรรค์.* พิมพ์ครั้งที่ 7. กรุงเทพฯ: ไทยวัฒนาพานิช.
- สร้อยญา จันทบุษกุล. (2561). *การวิเคราะห์ห่อภิมานงานวิจัยด้านนวัตกรรมการพัฒนาความคิดสร้างสรรค์ของนักเรียนระดับประถมศึกษา.* นครปฐม: คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร
- สุชาติดา ยอดสุรงค์. (2552). *การพัฒนาแบบวัดความคิดสร้างสรรค์ทักษะภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 1 สังกัดสำนักงานเขตพื้นที่การศึกษาขอนแก่น เขต 1.* (วิทยานิพนธ์ปริญญาโทบริหารศึกษิต). มหาวิทยาลัยราชภัฏเลย, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
- สุวิมล เสวกสุริยวงศ์. (2553). *การสังเคราะห์งานวิจัยเกี่ยวกับการสร้างแบบสอบวินิจฉัยทางคณิตศาสตร์.* (วิทยานิพนธ์ปริญญาโทบริหารศึกษิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
- สุวิมล ตีรกานันท์. (2551). *การสร้างเครื่องมือวัดตัวแปรในการวิจัยทางสังคมศาสตร์: แนวทางสู่การปฏิบัติ.* พิมพ์ครั้งที่ 2. กรุงเทพมหานคร: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สำนักงานเลขาธิการสภาการศึกษา. (2560). *แผนการศึกษาแห่งชาติ พ.ศ. 2560-2579.* กรุงเทพมหานคร: บริษัท พริกหวานกราฟฟิค จำกัด.
- สำนักงานเลขาธิการสภาการศึกษา. (2552). *รายงานการสังเคราะห์งานวิจัยเกี่ยวกับคุณภาพการศึกษาไทย: การวิเคราะห์ห่อภิมาน (Meta-analysis).* กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- สำนักวิชาการและมาตรฐานการศึกษา. (2557). *แนวปฏิบัติการวัดและประเมินผลการเรียนรู้ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551.* พิมพ์ครั้งที่ 4. กรุงเทพมหานคร: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด.
- อนุวัติ คุณแก้ว. (2558). *การวัดและประเมินผลการศึกษาแนวใหม่.* พิมพ์ครั้งที่ 2. กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.
- อภาพรรณ ประทุมไทย. (2553). *การวิเคราะห์ห่อภิมานคุณภาพของแบบวัดความถนัด.* (วิทยานิพนธ์ปริญญาโทบริหารศึกษิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
- อุทุมพร จามรมาน. (2531). *การสังเคราะห์งานวิจัย: เชิงปริมาณ.* กรุงเทพมหานคร: ฟันนี้พับบลิชซิง.
- Cheng, K-M. (2017). *Advancing 21st century competencies in East Asian education systems.* [Online]. Available from: <https://asiasociety.org/files/21st-century-competencies-east-asian-education-systems.pdf>[2019, Jun 11]

- Cheng, K-M., Jackson, L. & Lee, W. (2017). *Advancing 21st century competencies in Hong Kong*. [Online]. Available from: <https://asiasociety.org/files/21st-century-competencies-hong-kong.pdf>[2019, Jun 11]
- Cooper, H. (2017). *Research synthesis and meta-analysis : a step-by-step approach*. 5th ed. Los Angeles, CA : Sage.
- Greenstein, L. (2012). *Assessing 21st century skills: A guide to evaluating mastery and authentic learning*. Thousand Oaks, CA, US: Corwin Press.
- Gronlund, N. & Waugh, N. (2009). *Assessment of student achievement*. 9th ed. Upper Saddle River, NJ : Pearson
- Issacs, T., Zara, C. and Herbert, G. with Coombs, S. and Smith, C. (2013). *Key concept in educational assessment*. Thousand Oaks. CA: SAGE.
- Koricheva, J., Gurevitch, J., & Mengersen, K. (2013). *Handbook of meta-analysis in ecology and evolution*. New Jersey: Princeton University Press.
- McIntire, S. A., & Miller, L. A. (2000). *Foundations of psychological testing*. New York: McGraw-Hill.
- Metiri Group. (2003). *enGauge 21st century skills for 21st century learners*. [Online]. Available from :https://www.cwasd.k12.wi.us/highschl/newsfile1062_1.pdf[2019, Jun 11]
- Miller, D., Linn, R & Gronlund, N. (2009). *Measurement and Assessment in Teaching*. 10th ed. Upper Saddle River, NJ : Pearson
- Ministry of Education, Singapore. (2018). *21st century competencies*. [Online]. Available from: <https://www.moe.gov.sg/education/education-system/21st-century-competencies>[2019, Jun 11]
- Nitko, A. J. & Brookhart, S. M. (2007). *Educational assessment of students*. 5th ed. Upper Saddle NJ: Pearson Education.
- Partnership for 21st Century Learning. (2019). *Framework for 21st century learning definitions*. [Online]. Available from: http://static.battelleforkids.org/documents/p21/P21_Framework_DefinitionsBFK.pdf[2019, Jun 11]
- Reynolds, C., Livingston, R., and Willson, V. (2009). *Measurement and assessment in education*. 2nd ed. Boston: Pearson.
- World Health Organization. (1997). *Life Skills Education for Children and Adolescents in School*. Geneva: World Health Organization.

ภาคผนวก ก
รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์

รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
1	กชพรรณ เกสัชชา. (2560). การสร้างแบบวัดความคิดสร้างสรรค์ทางวิทยาศาสตร์สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
2	กติกกร กมลรัตน์สมบัติ. (2558). ผลของการให้ข้อมูลย้อนกลับจากแบบสอบถามย่อยที่มีต่อพัฒนาการความสามารถในการแก้ปัญหาทางคณิตศาสตร์ของนักเรียนระดับมัธยมศึกษา. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
3	กมลรัตน์ วัชรินทร์. (2547). การพัฒนาแบบวัดทักษะชีวิตสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1-3. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยเกษตรศาสตร์, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา.
4	กรรณา วิชระธำรงกุล. (2552). การสร้างแบบวัดความตระหนักรู้ต่อผลกระทบของสภาวะโลกร้อน สำหรับนักเรียนช่วงชั้นที่ 4 สังกัดสำนักงานเขตพื้นที่การศึกษา กรุงเทพมหานคร เขต 1. (ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, บัณฑิตวิทยาลัย, สาขาวิชาการวัดผลการศึกษา.
5	กฤษณา ปัญญา. (2552). การสร้างแบบวัดทักษะชีวิตของนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 6. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยเชียงใหม่, คณะศึกษาศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
6	กัญญ์พิชญา ยังดี. (2555). การสร้างแบบวัดความเข้าใจในการอ่านภาษาอังกฤษ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
7	กาญจนา สังข์ผาด. (2559). การพัฒนาแบบวัดความไว้วางใจระหว่างบุคคลในเครือข่ายสังคมออนไลน์สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลายในจังหวัดชลบุรี. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยบูรพา, คณะศึกษาศาสตร์, สาขาวิชาวิจัย วัดผลและสถิติการศึกษา.
8	เกียรติสุดา นาคูน. (2551). การพัฒนาแบบวัดคุณธรรม จริยธรรม และค่านิยมที่พึงประสงค์ สำหรับนักเรียนช่วงชั้นที่ 3-4. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏอุบลราชธานี, บัณฑิตวิทยาลัย, สาขาวิจัยและประเมินผลการศึกษา.
9	ขวัญยืน มูลศรี. (2548). การสร้างแบบวัดทักษะชีวิตสำหรับนักเรียนช่วงชั้นที่ 3 (ชั้นมัธยมศึกษาปีที่ 1-3). (ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, บัณฑิตวิทยาลัย, สาขาวิชาการวัดผลการศึกษา.
10	คำเพียร จันทน์แสน. (2553). การสร้างแบบวัดความสามารถการคิดอย่างมีวิจารณญาณสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏเพชรบูรณ์, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
11	จรงค์ษ์ ภาโส. (2553). การสร้างแบบวัดจิตวิทยาศาสตร์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏเพชรบูรณ์ คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
12	จริยาวดี ชวงศ์ศิริกุล. (2550). การพัฒนาแบบวัดทักษะกระบวนการทางคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษากงเก็ด. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยสุโขทัยธรรมาธิราช, สาขาวิชาศึกษาศาสตร์, แขนงวิชาการวัดและประเมินผลการศึกษา.
13	จันทร์จิรา อังกลสิทธิ์. (2552). การสร้างแบบวัดทักษะการคิดสำหรับนักเรียนช่วงชั้นที่ 3 เขตพื้นที่การศึกษาเชียงใหม่ เขต 1. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยเชียงใหม่, คณะศึกษาศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
14	จาดุพักตร์ พากเพียร. (2548). การศึกษาสรรพภาพทางสมองด้านการคิดวิเคราะห์และความคิดสร้างสรรค์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในจังหวัดศรีสะเกษ. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏอุบลราชธานี, บัณฑิตวิทยาลัย, สาขาวิจัยและประเมินผลการศึกษา.
15	จิราภรณ์ ศรีเมืองคุณ. (2558). การพัฒนาเครื่องมือวัดสมรรถนะผู้เรียนด้านเทคโนโลยีสารสนเทศ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษากาญจนบุรี เขต 4. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏกาญจนบุรี, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
16	จุฑามาศ ช่อคง. (2556). การสร้างแบบวัดความเป็นพลเมืองดีสำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 4-6 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 18. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยเกษตรศาสตร์, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา.
17	จุลพงษ์ กลิ่นหอม. (2549). การสร้างแบบวัดจิตวิทยาศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 5 กลุ่มโรงเรียนในเครือมูลนิธิเซนต์คาเบรียลแห่งประเทศไทย เขตกรุงเทพมหานครและปริมณฑล. (ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, บัณฑิตวิทยาลัย, สาขาวิชาการวัดผลการศึกษา.
18	เฉลิมศักดิ์ มะลิงาม. (2558). การพัฒนามาตรวัดจิตวิทยาศาสตร์สำหรับนักเรียนระดับมัธยมศึกษาตอนต้น : การประยุกต์ใช้การวิเคราะห์กลุ่มแฝงเพื่อกำหนดคะแนนจุดตัด. (วิทยานิพนธ์ปริญญามหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
19	ชฎานิชฐ์ โสทธิโยธิน. (2559). การสร้างแบบวัดการรับรู้ความสามารถของตนเองด้านวิทยาศาสตร์และคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษาเขต 2. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยเกษตรศาสตร์, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา.

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
20	ชนิดดา เทียนฤกษ์. (2557). การพัฒนาโมเดลการวัดทักษะชีวิตและอาชีพของนักเรียนระดับชั้นมัธยมศึกษาตอนปลายในศตวรรษที่ 21. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาวิธีวิทยาการวิจัยการศึกษา.
21	ชนิสรา สงวนไว้. (2558). การเปรียบเทียบความเที่ยงของแบบสอบวัดความสามารถในการแก้ปัญหาอย่างสร้างสรรค์ทางคณิตศาสตร์: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาวิธีวิทยาการวิจัยการศึกษา.
22	ชัยวิชิต เขียรชนะ. (2548). การพัฒนาแบบวัดความมีวินัยในตนเองของนักเรียนช่วงชั้นที่สาม : กรณีศึกษาจังหวัดนครนายก. (ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยศรีนครินทรวิโรฒ, บัณฑิตวิทยาลัย, สาขาวิชาการศึกษาและการวัดผลการศึกษา.
23	ดุสิต ทิบุญมา. (2555). การสร้างแบบวัดทักษะชีวิตสำหรับนักเรียนในระดับชั้นมัธยมศึกษาปีที่ 6. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
24	เทวินทร์ พิศวง. (2547). การพัฒนาแบบวัดคุณธรรม จริยธรรมและค่านิยมที่พึงประสงค์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จังหวัดจันทบุรี. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยสุโขทัยธรรมาธิราช, สาขาวิชาศึกษาศาสตร์, แขนงวิชาการศึกษาและการวัดผลการศึกษา.
25	ธวัชรัตน์ สีหานาจ. (2557). การพัฒนาแบบวัดจิตวิทยาศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 สำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 8 จังหวัดกาญจนบุรี. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏกาญจนบุรี, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
26	ธรรรัตน์ เกียรติวิบูลย์จิตร. (2560). การพัฒนาแบบวัดจิตวิทยาศาสตร์เชิงสถานการณ์สำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏนครราชสีมา, สำนักวิทยบริการและเทคโนโลยีสารสนเทศ., สาขาวิจัยและประเมินผลการศึกษา.
27	ธิดารัตน์ แตะเส็งซ้าย. (2548). การพัฒนาแบบวัดคุณลักษณะด้านจริยธรรมตามหลักสูตรการศึกษาขั้นพื้นฐานสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยสงขลานครินทร์, คณะศึกษาศาสตร์, สาขาวิชาการศึกษาและการวัดผลและการวิจัยการศึกษา.
28	นันทิดา ราศรี. (2553). การสร้างแบบสอบวัดการคิดอย่างมีวิจารณญาณเกี่ยวกับเศรษฐศาสตร์ตามแนวคิด เศรษฐกิจพอเพียงสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยขอนแก่น, คณะศึกษาศาสตร์, สาขาวิชาการศึกษาและการวัดผลการศึกษา.
29	น้ำทิพย์ งามอานาณิชัย. (2556). การพัฒนาแบบวัดทักษะแห่งศตวรรษที่ ๒๑ ตามการรับรู้ของผู้เรียนมัธยมศึกษาตอนต้น: การประยุกต์ใช้แนวคิดการเข้าถึงคุณลักษณะที่มุ่งวัดของ

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
	แบบสอบ. (วิทยานิพนธ์ปริญญาตรีบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
30	นิตารวรรณ บุญอินทร์. (2558). การพัฒนาแบบวัดความตระหนักเกี่ยวกับภาวะโลกร้อนของนักเรียนชั้นมัธยมศึกษาตอนต้น สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาอุทัยธานี เขต 1. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยเกษตรศาสตร์, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา.
31	นุตาพัทธ์ ลาภเกรียงไกร. (2557). การพัฒนาแบบวัดทักษะชีวิตของนักเรียนชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 1 (กรุงเทพมหานคร). (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยเกษตรศาสตร์, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา.
32	บุญสม ศรีศักดิ์. (2558). การพัฒนาเครื่องมือประเมินสมรรถนะสำคัญของนักเรียนมัธยมศึกษาตอนต้นตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551. (วิทยานิพนธ์ปริญญาตรีบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
33	บุศรา ชานูญมี. (2555). การสร้างแบบวัดความสามารถในการอ่านคิดวิเคราะห์และเขียนสื่อความของนักเรียนชั้นมัธยมศึกษาปีที่ 3. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิชาวิจัยและประเมินผลการศึกษา.
34	เบญจมาศ ณรงค์กุล. (2548). การพัฒนาเครื่องมือวัดคุณลักษณะอันพึงประสงค์ของนักเรียนช่วงชั้นที่ 3 ในโรงเรียนขยายโอกาสทางการศึกษา จังหวัดยะลา. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยทักษิณ, สาขาวิชาการวัดผลการศึกษา.
35	ปกรณ์ ประจัญบาน. (2559). การวิจัยและพัฒนาแบบวัดคุณลักษณะการเป็นบุคคลแห่งการเรียนรู้ ของนักเรียนชั้นมัธยมศึกษาในศตวรรษที่ 21.[ออนไลน์]. แหล่งที่มา: https://www.niets.or.th/uploads/content_pdf/pdf_1501733647.pdf [21 กันยายน 2562]
36	ปกรณ์ ประจัญบาน และ อนุชา กอนพวง. (2558). การวิจัยและพัฒนาแบบวัดทักษะในศตวรรษที่ 21 ด้านการรู้เท่าทันสื่อของนักเรียนชั้นมัธยมศึกษา.[ออนไลน์]. แหล่งที่มา: http://www.niets.or.th/uploads/content_pdf/pdf_1453439549.pdf [21 กันยายน 2562]
37	ประนอม บัวแก้ว. (2560). การพัฒนาชุดแบบวัดทักษะกระบวนการทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 40. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยสุโขทัยธรรมาธิราช, สาขาวิชาศึกษาศาสตร์, แขนงวิชาการวัดและประเมินผลการศึกษา.
38	ประภาส เทพแก้ว. (2554). การพัฒนาแบบวัดคุณลักษณะของผู้เรียนระดับมัธยมศึกษาตามแนวคิดการจัดการศึกษาเชิงคุณลักษณะ. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยเชียงใหม่, คณะศึกษาศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
39	ประยูรภรณ์ บุ่งทอง. (2555). การสร้างแบบวัดความมีวินัยในตนเองสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้นของโรงเรียนขยายโอกาสทางการศึกษา สังกัดสำนักงานเขตพื้นที่การศึกษาประถมศึกษาบุรีรัมย์ เขต 4. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
40	ปวีณา มะแซ. (2561). การพัฒนาแบบวัดทักษะการรู้เท่าทันสื่อในศตวรรษที่ 21 ของนักเรียนชั้นมัธยมศึกษาตอนต้นโดยประยุกต์ใช้ทฤษฎีการตอบสนองข้อสอบแบบพหุวิภาค. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยสงขลานครินทร์, คณะศึกษาศาสตร์, สาขาวิชาการวิจัยและประเมินผลการศึกษา.
41	ปาริชาติ อังกาบ. (2561). การสร้างแบบวัดทักษะชีวิตและอาชีพตามแนวคิดทักษะในศตวรรษที่ 21 สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
42	พรรณวดี ยืนยงค์นาน. (2554). การพัฒนาแบบวัดความตระหนักต่อโลกของนักเรียนมัธยมศึกษา : การทดสอบความไม่แปรเปลี่ยนของการวัดตามตัวแปรเพศ. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาวิธีวิทยาการวิจัยการศึกษา.
43	พัชรนันท์ มาอยู่วัง. (2552). การสร้างแบบวัดทักษะชีวิตสำหรับนักเรียนระดับช่วงชั้นที่ 3 สังกัดสำนักงานเขตพื้นที่การศึกษาเพชรบูรณ์ เขต 1. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏเพชรบูรณ์ คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
44	พัชรา คล่องขยัน. (2555). การสร้างแบบวัดจิตวิทยาศาสตร์แบบกำหนดสถานการณ์ของนักเรียนชั้นมัธยมศึกษาปีที่ 1. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
45	พัชริน สุภารี. (2550). การสร้างแบบวัดการคิดอย่างมีวิจารณญาณสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของเขตพื้นที่การศึกษาร้อยเอ็ดเขต 1. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยขอนแก่น, คณะศึกษาศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
46	พิไลลักษณ์ บัวทอง. (2554). การสร้างแบบวัดความสามารถในการแก้ปัญหาวิชาฟิสิกส์ เรื่องแรงและการเคลื่อนที่ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
47	ไพ สระแก้ว. (2555). การสร้างแบบวัดทักษะชีวิต สำหรับนักเรียนระดับชั้นมัธยมศึกษาปีที่ 3 สังกัดสำนักงานเขตพื้นที่การศึกษาร้อยเอ็ด เขต 2. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
48	ไพชยนต์ บุญสุภา. (2546). การพัฒนาแบบวัดความคิดสร้างสรรค์ด้านคณิตศาสตร์ เรื่องจำนวนและตัวเลข สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิชาการศึกษา.

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
	มหาบัณฑิต). มหาวิทยาลัยราชภัฏอุบลราชธานี, บัณฑิตวิทยาลัย, สาขาวิจัยและประเมินผลการศึกษา.
49	มนัญญา หาญอาสา. (2560). การพัฒนาแบบวัดพฤติกรรมการติดอินเทอร์เน็ตสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้นโรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 18. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยบูรพา, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยวัดผลและสถิติการศึกษา.
50	มาฆพันธ์ อำนาคิล. (2557). การเปรียบเทียบความเที่ยงของแบบสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาทางวิทยาศาสตร์ที่ใช้วิธีการตรวจให้คะแนนที่แตกต่างกันภายใต้จำนวนเหตุการณ์ที่ต่างกัน: การประยุกต์ใช้ทฤษฎีการสรุปอ้างอิงความน่าเชื่อถือของผลการวัด. (วิทยานิพนธ์ปริญญามหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาวิธีวิทยาการวิจัยการศึกษา.
51	ยุพิน ฐิตะสาร. (2550). การพัฒนาแบบทดสอบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏอุบลราชธานี, บัณฑิตวิทยาลัย, สาขาวิจัยและประเมินผลการศึกษา.
52	ยุภา เกิดศิริ. (2558). การพัฒนาแบบวัดความมีวินัยในตนเองสำหรับนักเรียนช่วงชั้นที่ 3 ของโรงเรียนมัธยมศึกษาในสังกัดสำนักงานเขตพื้นที่การศึกษากาญจนบุรี. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏกาญจนบุรี, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
53	รัตนา แซ่อู่. (2548). การพัฒนาเครื่องมือวัดความสามารถทางด้านเทคโนโลยีสารสนเทศสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในจังหวัดลำปาง. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยทักษิณ, สาขาวิชาการวัดผลการศึกษา.
54	ลัดดาวัลย์ วิจิตรจินดา. (2551). การพัฒนาเครื่องมือวัดการปฏิบัติการเขียนเรื่องภาษาอังกฤษเชิงสร้างสรรค์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จังหวัดสงขลา. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยทักษิณ, สาขาวิชาการวัดผลการศึกษา.
55	วราภรณ์ ไบภักดี. (2547). การพัฒนาแบบวัดการคิดอย่างมีวิจารณญาณ สำหรับนักเรียนช่วงชั้นที่ 3. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยนเรศวร, บัณฑิตวิทยาลัย, สาขาวิชาการวัดผลการศึกษา.
56	วาสนา กองณรงค์. (2559). การพัฒนาแบบวัดการร่วมมือรวมพลังเชิงสถานการณ์ สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏนครราชสีมา, สำนักวิทยบริการและเทคโนโลยีสารสนเทศ., สาขาวิจัยและประเมินผลการศึกษา.
57	วาสนา กองณรงค์. (2560) การพัฒนาแบบวัดทักษะและกระบวนการทางคณิตศาสตร์ เรื่องอัตราส่วนและร้อยละ ชั้นมัธยมศึกษาปีที่ 2. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
58	วาสนา ไกรแก้ว. (2556). การสร้างแบบวัดทักษะการแก้ปัญหาคณิตศาสตร์ เรื่อง เซต สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 4. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
59	วิภาภรณ์ เอียดคง. (2550). การพัฒนาเครื่องมือวัดความสามารถพิเศษทางคณิตศาสตร์ของนักเรียน ชั้นมัธยมศึกษาปีที่ 3 ในจังหวัดสงขลา. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยทักษิณ, สาขาวิชาการวัดผลการศึกษา.
60	วีรญา นรารัตน์. (2560). การพัฒนาแบบวัดทักษะการคิดอย่างมีวิจารณญาณในศตวรรษที่ 21 ระดับมัธยมศึกษาตอนต้น. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏนครราชสีมา, สำนักวิทยบริการและเทคโนโลยีสารสนเทศ., สาขาวิจัยและประเมินผลการศึกษา.
61	วุฒิพงษ์ นาสะอ้าน. (2560). การพัฒนาแบบทดสอบวัดความสามารถในการคิดอย่างมีวิจารณญาณทางด้านวิทยาศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 5. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
62	ศรัณยา หนูเงิน. (2549). การพัฒนาแบบวัดความสามารถทางการคิดตามทฤษฎีปัญญาแห่งความสำเร็จของสเติร์นเบอร์เกอร์สำหรับนักเรียนช่วงชั้นที่ 3. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยบูรพา, บัณฑิตวิทยาลัย, สาขาเทคโนโลยีการวัดทางการศึกษา.
63	ศศิธร เวียงอินทร์. (2547). การพัฒนาแบบวัดความคิดสร้างสรรค์ทางคณิตศาสตร์ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ของโรงเรียนในสังกัด สำนักงานเขตพื้นที่การศึกษา ร้อยเอ็ด เขต 1. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
64	ศิริเดช สุขิวะ และคณะ. (2559). ทักษะทางปัญญาของนักเรียนระดับมัธยมศึกษา: การพัฒนาโมเดลและเครื่องมือวัดออนไลน์.[ออนไลน์]. แหล่งที่มา: https://www.niets.or.th/uploads/content_pdf/pdf_1512096158.pdf [21 กันยายน 2562]
65	สมชัย จันทริมา. (2549). การพัฒนาแบบวัดคุณลักษณะอันพึงประสงค์นักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามหาสารคาม เขต 1. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
66	สร้อยญา จันทร์ชูสกุล, อัจศรา ประเสริฐสิน และพินดา วราสุนันท์. (2560). การพัฒนามาตรวัดการรู้เท่าทันสื่อและสารสนเทศสำหรับนักเรียนระดับชั้นมัธยมศึกษา.[ออนไลน์]. แหล่งที่มา: https://www.niets.or.th/uploads/content_pdf/pdf_1525928645.pdf [21 กันยายน 2562]
67	สังวาลย์ โพนพุท. (2556). การพัฒนาแบบวัดทักษะการอ่านจับใจความสำคัญวิชาภาษาอังกฤษ สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 1 ในโรงเรียนสังกัดสำนักงานเขตพื้นที่การศึกษามัธยมศึกษา เขต 21. (วิทยานิพนธ์ปริญญามหาบัณฑิต).

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
	มหาวิทยาลัยสุโขทัยธรรมมาธิราช, สาขาวิชาศึกษาศาสตร์, แขนงวิชาการวัดและประเมินผลการศึกษา.
68	สาคร สียงนอก. (2556). การพัฒนาแบบวัดทักษะกระบวนการทางคณิตศาสตร์สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
69	สุชาติ ยอดสุรางค์. (2552). การพัฒนาแบบวัดความคิดสร้างสรรค์ทักษะภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 1 สังกัดสำนักงานเขตพื้นที่การศึกษาขอนแก่น เขต 1. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏเลย, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
70	สุชาติ ใจสถาน. (2553). การพัฒนาแบบวัดจริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสารสำหรับนักเรียน. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยบูรพา, บัณฑิตวิทยาลัย, สาขาวิชาวิจัย วัดผลและสถิติการศึกษา.
71	สุทธิวรรณ พิรศักดิ์โสภณ, อรณัฐ ศรีสะอาด, รุ่งฤดี กล้าหาญ และพนัส จันทร์เปล่ง. (2559). การพัฒนาเครื่องมือวัดและประเมินคุณลักษณะอันพึงประสงค์รอบด้าน ตามมาตรฐานการศึกษาแห่งชาติ: มิติด้านจิตใจ ของนักเรียนระดับชั้นประถมศึกษาปีที่ 6 มัธยมศึกษาปีที่ 3 และมัธยมศึกษาปีที่ 6. [ออนไลน์]. แหล่งที่มา: https://www.niets.or.th/uploads/content_pdf/pdf_1512096010.pdf [21 กันยายน 2562]
72	สุนารี มีใหม่. (2557). การพัฒนาแบบวัดจิตวิทยาศาสตร์ของนักเรียนมัธยมศึกษาตอนปลาย: การวิเคราะห์ความไม่แปรเปลี่ยนของโมเดลการวัดระหว่างแผนการเรียน. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาวิธีวิทยาการวิจัยการศึกษา.
73	สุพรรณิ ดวงแก้ว. (2555). การพัฒนาแบบวัดความคิดสร้างสรรค์ทางศิลปะ (ทัศนศิลป์) สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏอุบลราชธานี, บัณฑิตวิทยาลัย, สาขาวิจัยและประเมินผลการศึกษา.
74	สุพัชญา เจริรัตน์. (2560). การเปรียบเทียบคุณสมบัติทางจิตมิติและประสิทธิภาพของแบบวัดสถานการณ์การรู้สื่อดิจิทัลที่มีรูปแบบต่างกัน. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการศึกษาและประเมินผลการศึกษา.
75	สุพัฒตา ภูสอดสี. (2553). การสร้างแบบทดสอบเอ็มไอวัดความสามารถในการแก้ปัญหาทางคณิตศาสตร์เรื่องเรขาคณิตวิเคราะห์ ชั้นมัธยมศึกษาปีที่ 4. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
76	สุพัตรา สงกลิ่น. (2551). การสร้างแบบวัดความสามารถในการอ่าน คติวิเคราะห์และเขียนสื่อความ สำหรับนักเรียนเตรียมทหาร โรงเรียนเตรียมทหาร. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยเกษตรศาสตร์, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา.

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
77	สุพัตรา แสงสุวรรณ. (2549). การพัฒนาแบบวัดความสามารถการคิด สำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
78	สุพิศพรรณ ภาระพันธ์. (2555). การสร้างแบบวัดการอ่านเชิงคิดวิเคราะห์จากบทอ่านภาษาอังกฤษสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
79	สุภาพร จันทร์ดอกไม้. (2553). การพัฒนาแบบวัดทักษะการคิดขั้นสูงสำหรับนักเรียนมัธยมศึกษาตอนต้น. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
80	สุมาลี คำชูรุ. (2555). การสร้างแบบวัดความสามารถในการอ่านคิดวิเคราะห์และเขียนสื่อความกลุ่มสาระการเรียนรู้ภาษาไทย สำหรับชั้นมัธยมศึกษาปีที่ 2. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
81	สุไรยา หมัดหมั่น. (2549). การพัฒนาแบบทดสอบอัตนัยประยุกต์วัดความสามารถในการแก้ปัญหาวทางวิทยาศาสตร์ ของนักเรียนชั้นมัธยมศึกษาปีที่ 5. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยทักษิณ, สาขาวิชาการวัดผลการศึกษา.
82	สุวรรณ ใจกล้า. 2561. การพัฒนาชุดเครื่องมือประเมินทักษะการเรียนรู้และนวัตกรรมตามแนวคิดการประเมินที่ใช้การปฏิบัติเป็นฐาน สำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยขอนแก่น, บัณฑิตวิทยาลัย, สาขาวิชาการวัดและประเมินผลการศึกษา.
83	หทัยชนก กุระสุวรรณ. (2560). การพัฒนาแบบวัดทักษะการเขียนเชิงวิเคราะห์สำหรับนักเรียนมัธยมศึกษาตอนต้น. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
84	หนูเวียง อนุภักดิ์. การสร้างแบบสอบการคิดอย่างมีวิจารณญาณเกี่ยวกับสารเสพติดสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 2. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยขอนแก่น, คณะศึกษาศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
85	อนันดา สันฐิตวิณิชัย. (2551). การพัฒนาแบบวัดความสามารถในการอ่าน คิดวิเคราะห์ และเขียนสื่อความสำหรับนักเรียนชั้นมัธยมศึกษาปีที่ 3. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). จุฬาลงกรณ์มหาวิทยาลัย, คณะครุศาสตร์, สาขาวิชาการวัดและประเมินผลการศึกษา.
86	ออมสิน ศรีปลั่ง. (2549). การพัฒนาแบบวัดคุณลักษณะอันพึงประสงค์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 โรงเรียนในสังกัดสำนักงานเขตพื้นที่การศึกษาร้อยเอ็ด เขต 3. (วิทยานิพนธ์ปริญญาโทมหาบัณฑิต). มหาวิทยาลัยมหาสารคาม, คณะศึกษาศาสตร์, สาขาวิชาการวัดผลการศึกษา.
87	อัจตรา ประเสริฐสิน, อนันดา สันฐิตวิณิชัย, ยური ผลพันธิน และมิ่งขวัญ ภาคสัญไชย. (2559). การพัฒนาตัวบ่งชี้และแบบวัดคุณลักษณะอันพึงประสงค์รอบด้านในด้านปัจจัยความสามารถ และสมรรถนะทางจิตใจของนักเรียนระดับชั้นประถมศึกษาและ

เรื่องที่	รายการอ้างอิงงานวิจัยที่ใช้ในการสังเคราะห์
	มัธยมศึกษา. กรุงเทพมหานคร. สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน).
88	อัมพร ปานจันดี. (2556). การพัฒนาแบบวัดความสามารถในการคิดอย่างมีวิจารณญาณสำหรับนักเรียนชั้นมัธยมศึกษาตอนต้น. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏมหาสารคาม, คณะครุศาสตร์, สาขาวิจัยและประเมินผลการศึกษา.
89	อำนาจ หุ่มนริน. (2558). การสร้างแบบวัดทักษะชีวิตสำหรับนักเรียนชั้นมัธยมศึกษาตอนปลาย. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยเกษตรศาสตร์, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา.
90	อำมาลา สารชาติ. (2549). การพัฒนาแบบทดสอบ เอ็ม อี คิว สำหรับวัดความสามารถ ในการแก้โจทย์ปัญหาคณิตศาสตร์ของนักเรียนระดับชั้นมัธยมศึกษาปีที่ 1. (วิทยานิพนธ์ปริญญามหาบัณฑิต). มหาวิทยาลัยราชภัฏอุบลราชธานี, บัณฑิตวิทยาลัย, สาขาวิจัยและประเมินผลการศึกษา.
91	อุไร จักษ์ตรีมงคล. (2556). การพัฒนาแบบวัดทักษะการคิดสำหรับการประเมินคุณภาพผู้เรียนระดับชั้นมัธยมศึกษาปีที่ 6. กรุงเทพมหานคร.[ออนไลน์]. แหล่งที่มา: https://www.niets.or.th/uploads/content_pdf/research_1391414572.pdf [21 กันยายน 2562]
92	พิชญา ตีมี. (2559). การพัฒนาแนวทางการประเมินทักษะการเรียนรู้และนวัตกรรมของผู้เรียนตามแนวคิดการประเมินผู้เรียนในศตวรรษที่ 21. (วิทยานิพนธ์ปริญญาคุษฎ์บัณฑิต). มหาวิทยาลัยนเรศวร, บัณฑิตวิทยาลัย, สาขาวิชาวิจัยและประเมินผลการศึกษา

ภาคผนวก ข
รายชื่อผู้เชี่ยวชาญ

ผู้เชี่ยวชาญในการตรวจสอบคุณภาพเครื่องมือ

รายชื่อ	ความเชี่ยวชาญ	สังกัด
1. รศ.ดร.กมลวรรณ ตังชันกานนท์	การวิจัย, การวัดประเมินผล	คณะครุศาสตร์
2. รศ.ดร.ณัฐภรณ์ หลาวทอง	การวิจัย, การวัดประเมินผล	จุฬาลงกรณ์มหาวิทยาลัย
3. รศ.ดร.ไชยยศ ไพวิทยศิริธรรม	การวิจัย, การวัดประเมินผล, การวิเคราะห์ห่อถัก	คณะศึกษาศาสตร์ มหาวิทยาลัยศิลปากร
4. ผศ.ดร.พินดา วราสุนันท์	การวิจัย, การวัดประเมินผล	คณะศึกษาศาสตร์และพัฒนศาสตร์ มหาวิทยาลัยเกษตรศาสตร์
5. ผศ.ดร.วราพร เอราวรรณ	การวิจัย, การวัดประเมินผล	คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

ภาคผนวก ค
เครื่องมือที่ใช้ในการวิจัย

แบบประเมินคุณภาพงานวิจัย

ชื่อผู้วิจัย.....

ชื่องานวิจัย.....

คำชี้แจง โปรดทำเครื่องหมาย ✓ ลงในช่องที่ตรงกับลักษณะงานวิจัยที่ประเมิน

ข้อ	ลักษณะงานวิจัยที่ประเมิน	ผลการประเมิน				
		0	1	2	3	4
1	ชื่อเรื่องมีความชัดเจน น่าสนใจ	0	1	2	3	4
2	ความเป็นมาและความสำคัญของปัญหา ชัดเจน สอดคล้องกับเรื่องที่ทำวิจัย	0	1	2	3	4
3	ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง ถูกต้องตามหลักการวิจัย	0	1	2	3	4
4	เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล	0	1	2	3	4
5	กำหนดขอบเขตของการวิจัยกำหนดได้อย่างเหมาะสม และมีเหตุผลรองรับ	0	1	2	3	4
6	การนิยามศัพท์เฉพาะมีความชัดเจน	0	1	2	3	4
7	กรอบแนวคิดในการวิจัยมีความถูกต้องเหมาะสม ชัดเจน ตามหลักการวิจัย	0	1	2	3	4
8	เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ	0	1	2	3	4
9	เอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย	0	1	2	3	4
10	เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัย	0	1	2	3	4
11	เอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศมีสัดส่วนที่เหมาะสม	0	1	2	3	4
12	การออกแบบการวิจัยสอดคล้องกับปัญหาการวิจัย	0	1	2	3	4
13	ขั้นตอนการวิจัยมีความชัดเจน	0	1	2	3	4
14	กำหนดประชากรและตัวอย่างได้อย่างเหมาะสมกับการวิจัย	0	1	2	3	4
15	ขนาดตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม	0	1	2	3	4
16	การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน	0	1	2	3	4
17	การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีแนวคิด ทฤษฎีเกี่ยวกับตัวแปรที่วัดและวิธีการสร้างเครื่องมือ	0	1	2	3	4
18	การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา	0	1	2	3	4
19	เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ที่สร้างมีความเป็นปรนัยของข้อคำถาม	0	1	2	3	4
20	ขั้นตอนในการดำเนินการวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	0	1	2	3	4
21	วิธีการตรวจสอบคุณภาพเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	0	1	2	3	4
22	มีการทดลองใช้เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 (try out)	0	1	2	3	4
23	ตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม	0	1	2	3	4
24	เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ในภาพรวมมีคุณภาพเหมาะสม	0	1	2	3	4
25	กระบวนการเก็บรวบรวมข้อมูลมีความชัดเจนและเหมาะสม	0	1	2	3	4
26	สถิติที่ใช้ในการวิเคราะห์ข้อมูลมีความถูกต้องเหมาะสมกับลักษณะของข้อมูล	0	1	2	3	4
27	ลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูล	0	1	2	3	4

ข้อ	ลักษณะงานวิจัยที่ประเมิน	ผลการประเมิน				
		0	1	2	3	4
28	การแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องชัดเจน	0	1	2	3	4
29	สรุปผลการวิจัยอย่างถูกต้อง	0	1	2	3	4
30	มีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีต	0	1	2	3	4
31	ข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์	0	1	2	3	4
32	เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางปฏิบัติ	0	1	2	3	4
33	เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ	0	1	2	3	4
34	รูปแบบรายงานถูกต้องตามหลักวิชา	0	1	2	3	4

เกณฑ์การประเมินคุณภาพสำหรับแบบประเมินคุณภาพงานวิจัย

1. ชื่อเรื่องมีความชัดเจน น่าสนใจ
 - 0 หมายถึง ไม่มีความชัดเจนหรือไม่ระบุถึงปัญหาวิจัย
 - 1 หมายถึง ระบุถึงปัญหาวิจัย
 - 2 หมายถึง ระบุถึงปัญหาวิจัย ระบุตัวแปรสำคัญที่ศึกษาแต่ไม่ครบถ้วน
 - 3 หมายถึง ระบุถึงปัญหาวิจัย ระบุตัวแปรสำคัญที่ศึกษาครบถ้วน
 - 4 หมายถึง ระบุถึงปัญหาวิจัย ระบุตัวแปรสำคัญที่ศึกษาครบถ้วน และระบุตัวอย่างที่ศึกษา
2. ความเป็นมาและความสำคัญของปัญหา ชัดเจน สอดคล้องกับเรื่องที่ทำวิจัย
 - 0 หมายถึง ไม่ระบุความเป็นมาและความสำคัญของปัญหา
 - 1 หมายถึง ความเป็นมาและความสำคัญของปัญหาไม่เกี่ยวข้องกับเรื่องที่ทำวิจัย ไม่ได้ระบุให้เห็นประเด็นของปัญหาและความสำคัญในการทำวิจัย
 - 2 หมายถึง ความเป็นมาและความสำคัญของปัญหาแสดงให้เห็นประเด็นปัญหาและความสำคัญในการทำวิจัย แต่ยังไม่ครอบคลุมประเด็นที่ศึกษา เนื้อหาไม่ต่อเนื่องสอดคล้องกัน
 - 3 หมายถึง ความเป็นมาและความสำคัญของปัญหาแสดงให้เห็นประเด็นปัญหาและความสำคัญในการทำวิจัย ครอบคลุมประเด็นที่ศึกษาโดยมีเหตุผลสนับสนุน เนื้อหาไม่ต่อเนื่องสอดคล้อง
 - 4 หมายถึง ความเป็นมาและความสำคัญของปัญหาแสดงให้เห็นประเด็นปัญหาและความสำคัญในการทำวิจัย ครอบคลุมประเด็นที่ศึกษาโดยมีเหตุผลสนับสนุน ข้อความมีความกระชับตรงประเด็น
3. ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง ถูกต้องตามหลักการวิจัย
 - 0 หมายถึง ปัญหาวิจัย/วัตถุประสงค์ไม่มีความสอดคล้องกับชื่อเรื่อง
 - 1 หมายถึง ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง
 - 2 หมายถึง ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง ระบุตัวแปรสำคัญที่ศึกษา
 - 3 หมายถึง ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง ระบุตัวแปรสำคัญที่ศึกษา ระบุวิธีที่ศึกษา
 - 4 หมายถึง ปัญหาวิจัย/วัตถุประสงค์มีความสอดคล้องกับชื่อเรื่อง ระบุตัวอย่างที่ศึกษา ระบุตัวแปรสำคัญที่ศึกษา ระบุวิธีที่ศึกษา
4. เหตุผลและความจำเป็นในการทำวิจัยมีความสมเหตุสมผล
 - 0 หมายถึง เป็นเรื่องที่ไม่มีความจำเป็นในการทำวิจัย
 - 1 หมายถึง เป็นเรื่องที่น่าสนใจเพียงอย่างเดียว
 - 2 หมายถึง เป็นเรื่องที่ต้องรู้ใหม่ และไม่มีผู้ใดเคยทำมาก่อน
 - 3 หมายถึง เป็นเรื่องที่น่าสนใจ ต้องรู้ใหม่ และไม่มีผู้ใดเคยทำมาก่อน
 - 4 หมายถึง เป็นเรื่องที่น่าสนใจ ต้องรู้ใหม่ ไม่เคยมีผู้ใดเคยทำมาก่อน และเป็นปัญหาเร่งด่วน
5. กำหนดขอบเขตของการวิจัยกำหนดได้อย่างเหมาะสม และมีเหตุผลรองรับ
 - 0 หมายถึง ไม่ระบุขอบเขตของการวิจัย
 - 1 หมายถึง ระบุขอบเขตของการวิจัยเพียง 1 ส่วน คือ กรอบความคิดตามทฤษฎี
 - 2 หมายถึง ระบุขอบเขตของการวิจัยครบ 2 ส่วน คือ กรอบความคิดตามทฤษฎี ระบุว่าตัวแปรในการวิจัยไม่ครบตามกรอบความคิด
 - 3 หมายถึง ระบุขอบเขตของการวิจัยครบ 3 ส่วน คือ กรอบความคิดตามทฤษฎี ระบุว่าตัวแปรในการวิจัยไม่ครบตามกรอบความคิด และอธิบายเหตุผลตัวแปรที่ไม่ครบกรอบความคิด
 - 4 หมายถึง ระบุขอบเขตของการวิจัยครบ 4 ส่วน คือ กรอบความคิดตามทฤษฎี ระบุว่าตัวแปรในการวิจัยไม่ครบตามกรอบความคิด และอธิบายเหตุผลตัวแปรที่ไม่ครบกรอบความคิด และระบุว่าผลการวิจัยสามารถ generalize ได้ตามกรอบความคิด
6. การนิยามศัพท์เฉพาะมีความชัดเจน
 - 0 หมายถึง ไม่มีการนิยามศัพท์เฉพาะ หรือมีการให้นิยามเชิงทฤษฎีแต่ไม่ครบถ้วน
 - 1 หมายถึง มีการนิยามตัวแปรสำคัญที่ศึกษาอย่างครบถ้วนแต่เป็นนิยามเชิงทฤษฎี

- 2 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัดแต่ไม่ครบทุกตัวแปรสำคัญที่ศึกษา
- 3 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัดครบทุกตัวแปรสำคัญที่ศึกษา
- 4 หมายถึง มีการให้นิยามเชิงทฤษฎี นิยามเชิงปฏิบัติการขององค์ประกอบที่ต้องการวัดครบทุกตัวแปรสำคัญที่ศึกษา โดยบ่งบอกลักษณะที่สำคัญในรูปที่สามารถสังเกตและวัดได้
7. กรอบแนวคิดในการวิจัยมีความถูกต้องเหมาะสม ชัดเจน ตามหลักการวิจัย
- 0 หมายถึง ไม่ระบุกรอบแนวคิด ไม่แสดงทฤษฎีหรือหลักฐานที่มาของกรอบแนวคิด
- 1 หมายถึง แสดงทฤษฎีหรือหลักฐานที่มาของกรอบแนวคิดบางส่วน ไม่มีการสังเคราะห์สร้างกรอบแนวคิดให้เห็น
- 2 หมายถึง ระบุกรอบแนวคิดโดยแสดงทฤษฎีหรือหลักฐานที่มาของกรอบแนวคิดนั้น แต่ไม่มีการสังเคราะห์สร้างกรอบแนวคิดให้เห็น
- 3 หมายถึง ระบุกรอบแนวคิดโดยแสดงทฤษฎีหรือหลักฐานที่มาของกรอบแนวคิดนั้น มีการสังเคราะห์ สรุปกรอบแนวคิด แต่ไม่มีแผนภาพความสัมพันธ์ระหว่างตัวแปรหรือแสดงแผนภาพความสัมพันธ์ระหว่างตัวแปรแต่ไม่ถูกต้อง
- 4 หมายถึง ระบุกรอบแนวคิดโดยแสดงทฤษฎีหรือหลักฐานที่มาของกรอบแนวคิดนั้น มีการสังเคราะห์ สรุปกรอบแนวคิด แสดงแผนภาพแสดงความสัมพันธ์ระหว่างตัวแปรอย่างถูกต้อง
8. เอกสารและงานวิจัยที่เกี่ยวข้องมีปริมาณมากพอ
- 0 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีจำนวนไม่ถึง 5 เล่ม
- 1 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีจำนวน 5 – 10 เล่ม
- 2 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีจำนวน 11 – 15 เล่ม
- 3 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีจำนวน 16 – 20 เล่ม
- 4 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีจำนวนมากกว่า 20 เล่ม
9. เอกสารและงานวิจัยที่เกี่ยวข้องมีความสอดคล้องกับปัญหา/วัตถุประสงค์ของการวิจัย
- 0 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องไม่สอดคล้องกับเรื่องที่ทำวิจัย ใช้ประโยชน์ไม่ได้
- 1 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องสอดคล้องกับเรื่องที่ทำวิจัย แต่ให้รายละเอียดไม่ชัดเจนไม่มีการสังเคราะห์สรุปเนื้อหา
- 2 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องสอดคล้องกับเรื่องที่ทำวิจัย ให้รายละเอียดชัดเจน มีการสังเคราะห์สรุปเนื้อหาในบางส่วน
- 3 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องสอดคล้องกับเรื่องที่ทำวิจัย ให้รายละเอียดชัดเจน มีการสังเคราะห์สรุปเนื้อหาทุกหัวข้อ แต่ข้อความไม่กระชับ เยิ่นเย้อ อ่านเข้าใจยาก
- 4 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องสอดคล้องกับเรื่องที่ทำวิจัย ให้รายละเอียดชัดเจน มีการสังเคราะห์สรุปเนื้อหาทุกหัวข้อ ข้อความกระชับ อ่านเข้าใจง่าย
10. เอกสารและงานวิจัยที่เกี่ยวข้องมีความทันสมัยและเชื่อมโยงกับงานวิจัยในอดีต
- 0 หมายถึง ไม่มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปีเลย
- 1 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณไม่ถึง 30%
- 2 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณตั้งแต่ 30% แต่ไม่ถึง 50%
- 3 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณตั้งแต่ 50% แต่ไม่ถึง 70%
- 4 หมายถึง มีเอกสารและงานวิจัยที่เกี่ยวข้องที่มีอายุระหว่าง 1 – 5 ปี มีปริมาณตั้งแต่ 70% ขึ้นไป
11. เอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศมีสัดส่วนที่เหมาะสม
- 0 หมายถึง ไม่มีเอกสารและงานวิจัยที่เกี่ยวข้องทั้งภาษาไทยและภาษาต่างประเทศ
- 1 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีเพียงภาษาไทยเท่านั้น
- 2 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทยมากกว่า 80% และภาษาต่างประเทศน้อยกว่า 20%

- 3 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทย 50% และภาษาต่างประเทศ 50%
- 4 หมายถึง เอกสารและงานวิจัยที่เกี่ยวข้องมีสัดส่วนเป็นภาษาไทยน้อยกว่า 30% และภาษาต่างประเทศมากกว่า 70%
12. การออกแบบการวิจัยสอดคล้องกับปัญหาการวิจัย
- 0 หมายถึง การออกแบบการวิจัยไม่สอดคล้องกับปัญหาการวิจัย
- 1 หมายถึง การออกแบบการวิจัยทำได้แนวทางการวิจัยที่จะได้คำตอบตรงประเด็นกับปัญหาวิจัย
- 2 หมายถึง การออกแบบการวิจัยทำให้ได้ผลการวิจัยมีความตรงภายใน หรือความตรงภายนอกอย่างใดอย่างหนึ่ง
- 3 หมายถึง การออกแบบการวิจัยทำให้ได้ผลการวิจัยมีความตรงภายในและความตรงภายนอก
- 4 หมายถึง การออกแบบการวิจัยทำได้แนวทางการวิจัยที่จะได้คำตอบตรงประเด็นปัญหาวิจัย และทำให้ผลการวิจัยมีความตรงภายใน และความตรงภายนอก
13. ขั้นตอนการวิจัยมีความชัดเจน
- 0 หมายถึง ไม่มีการอธิบายขั้นตอนการวิจัย
- 1 หมายถึง อธิบายวิธีการดำเนินการวิจัยไม่ครบถ้วน
- 2 หมายถึง อธิบายวิธีการดำเนินการวิจัยเป็นขั้นตอน แต่ไม่ชัดเจนทุกขั้นตอน
- 3 หมายถึง อธิบายวิธีการดำเนินการวิจัยเป็นขั้นตอนอย่างเหมาะสม และชัดเจนทุกขั้นตอน
- 4 หมายถึง อธิบายวิธีการดำเนินการวิจัยเป็นขั้นตอนอย่างเหมาะสม และชัดเจนทุกขั้นตอน รวมทั้งมีแผนภูมิแสดงขั้นตอนการดำเนินงาน
14. กำหนดประชากรและตัวอย่างได้อย่างเหมาะสมกับการวิจัย
- 0 หมายถึง ไม่ระบุประชากร และตัวอย่าง
- 1 หมายถึง ระบุประชากร และตัวอย่างเพียงอย่างใดอย่างหนึ่ง
- 2 หมายถึง ระบุประชากร ระบุตัวอย่าง แต่ไม่มีการสุ่มตัวอย่าง
- 3 หมายถึง ระบุประชากร ระบุตัวอย่าง ระบุการสุ่มตัวอย่าง แต่ให้รายละเอียดในการสุ่มตัวอย่างไม่ชัดเจน
- 4 หมายถึง ระบุประชากร ระบุตัวอย่าง ระบุการสุ่มตัวอย่าง ให้รายละเอียดในการสุ่มตัวอย่างชัดเจน อ่านเข้าใจง่าย
15. ขนาดตัวอย่างและวิธีการสุ่มมีความถูกต้องเหมาะสม
- 0 หมายถึง ไม่มีการระบุการได้มาของตัวอย่างและขนาดตัวอย่าง หรือศึกษากับประชากร
- 1 หมายถึง ระบุขนาดตัวอย่าง แต่ไม่มีการระบุการได้มาของตัวอย่าง
- 2 หมายถึง ระบุขนาดตัวอย่าง และตัวอย่างได้มาจากการเลือกตัวอย่างโดยไม่อาศัยความน่าจะเป็น
- 3 หมายถึง ระบุขนาดตัวอย่าง และตัวอย่างได้มาจากการสุ่มโดยอาศัยความน่าจะเป็น และเป็นตัวแทนที่ดีของประชากร แต่มีที่มาของการกำหนดขนาดของตัวอย่างไม่สมเหตุผล
- 4 หมายถึง ระบุขนาดตัวอย่าง และตัวอย่างได้มาจากการสุ่มโดยอาศัยความน่าจะเป็น เป็นตัวแทนที่ดีของประชากร และมีที่มาของการกำหนดขนาดของตัวอย่างสมเหตุผล
16. การกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 มีความชัดเจน
- 0 หมายถึง ไม่มีการกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21
- 1 หมายถึง มีการกำหนดจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 แต่ไม่ชัดเจน
- 2 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน แต่ไม่สอดคล้องกับความต้องการนำไปใช้ประโยชน์
- 3 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน และสอดคล้องกับความต้องการนำไปใช้ประโยชน์ แต่ไม่ครบถ้วนในบางประเด็น
- 4 หมายถึง มีการระบุจุดมุ่งหมายในการวัดสมรรถนะในศตวรรษที่ 21 ที่ชัดเจน ครบถ้วน และสอดคล้องกับความต้องการของสังคมยุคปัจจุบันในการนำไปใช้ประโยชน์
17. การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีแนวคิด ทฤษฎีเกี่ยวกับตัวแปรที่วัดและวิธีการสร้างเครื่องมือ
- 0 หมายถึง ไม่มีการศึกษาแนวคิด ทฤษฎีเกี่ยวกับตัวแปรที่วัด รวมถึงวิธีการสร้างและหาคุณภาพเครื่องมือ

- 1 หมายถึง มีการศึกษาวิธีการสร้างและหาคุณภาพเครื่องมือ แต่ไม่มีการศึกษาแนวคิด ทฤษฎีเกี่ยวกับตัวแปรที่วัด
 - 2 หมายถึง มีการศึกษาแนวคิด ทฤษฎีเกี่ยวกับตัวแปรที่วัด และวิธีการสร้าง แต่ไม่มีการศึกษาการหาคุณภาพเครื่องมือ
 - 3 หมายถึง มีการศึกษาแนวคิด ทฤษฎีเกี่ยวกับตัวแปรที่วัด วิธีการสร้างและหาคุณภาพเครื่องมือ แต่ขาดความสมบูรณ์ในบางประเด็น
 - 4 หมายถึง มีการศึกษาแนวคิด ทฤษฎีเกี่ยวกับตัวแปรที่วัด วิธีการสร้างและหาคุณภาพเครื่องมืออย่างครบถ้วน
18. การสร้างเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีการกำหนดแผนผังการสร้างข้อคำถามที่ครอบคลุมเนื้อหา
- 0 หมายถึง ไม่มีการกำหนดแผนผังการสร้างข้อคำถาม
 - 1 หมายถึง มีการกำหนดแผนผังการสร้างข้อคำถาม แต่ไม่มีการแยกรายละเอียดตามองค์ประกอบ ตัวชี้วัด ที่ต้องการวัด
 - 2 หมายถึง มีการกำหนดแผนผังการสร้างข้อคำถาม และแยกรายละเอียดตามองค์ประกอบ ตัวชี้วัด ที่ต้องการวัด
 - 3 หมายถึง มีการกำหนดแผนผังการสร้างข้อคำถาม และแยกรายละเอียดตามองค์ประกอบ ตัวชี้วัด ที่ต้องการวัด และมีการระบุเลขข้อ จำนวนข้อที่วัดตามองค์ประกอบนั้น ๆ แต่ขาดความเหมาะสมในบางประเด็น
 - 4 หมายถึง มีการกำหนดแผนผังการสร้างข้อคำถาม และแยกรายละเอียดตามองค์ประกอบ ตัวชี้วัด ที่ต้องการวัด และมีการระบุเลขข้อ จำนวนข้อที่วัดตามองค์ประกอบนั้น ๆ ได้อย่างเหมาะสมและครอบคลุมเนื้อหา
19. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ที่สร้างมีความเป็นปรนัยของข้อคำถาม
- 0 หมายถึง ลักษณะคำถามส่วนใหญ่ไม่ชัดเจน คลุมเครือ และไม่มีกระบวนการตรวจสอบให้คะแนน/ไม่มีตัวอย่างเครื่องมือในรายงานการวิจัย
 - 1 หมายถึง ลักษณะคำถามส่วนใหญ่ไม่ชัดเจน คลุมเครือ แต่มีการกระบวนการตรวจสอบให้คะแนน
 - 2 หมายถึง ลักษณะคำถามส่วนใหญ่ชัดเจน เข้าใจง่าย แต่ไม่มีกระบวนการตรวจสอบให้คะแนน
 - 3 หมายถึง ลักษณะคำถามส่วนใหญ่ชัดเจน เข้าใจง่าย ไม่กำกวม และมีการกระบวนการตรวจสอบให้คะแนน
 - 4 หมายถึง ลักษณะคำถามทุกข้อชัดเจน เข้าใจง่าย ไม่กำกวม และมีการกระบวนการตรวจสอบให้คะแนนอย่างละเอียด
20. ขั้นตอนในการดำเนินการวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม
- 0 หมายถึง ขั้นตอนในการดำเนินการวัดไม่ชัดเจน ไม่มีการระบุค่าชี้แจงในเครื่องมือ
 - 1 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน แต่ไม่มีการระบุค่าชี้แจงในเครื่องมือ
 - 2 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียด แต่ไม่ครบถ้วน
 - 3 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียดครบถ้วน
 - 4 หมายถึง ขั้นตอนในการดำเนินการวัดชัดเจน มีการระบุค่าชี้แจงในเครื่องมือ และมีรายละเอียดครบถ้วน พร้อมยกตัวอย่างวิธีการทำ
21. วิธีการตรวจสอบคุณภาพเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม
- 0 หมายถึง ไม่มีการตรวจสอบคุณภาพเครื่องมือ
 - 1 หมายถึง มีการวิเคราะห์ค่าสถิติพื้นฐานเพียงอย่างเดียว
 - 2 หมายถึง มีการวิเคราะห์ค่าสถิติพื้นฐานและมีการวิเคราะห์ข้อคำถามรายข้อ (ความยาก อำนาจจำแนก)
 - 3 หมายถึง มีการวิเคราะห์ค่าสถิติพื้นฐาน การวิเคราะห์ข้อคำถามรายข้อ (ความยาก อำนาจจำแนก) และมีการวิเคราะห์ค่าความเที่ยง ความตรง แต่เลือกวิธีการไม่เหมาะสม
 - 4 หมายถึง มีการวิเคราะห์ค่าสถิติพื้นฐานและมีการวิเคราะห์ข้อคำถามรายข้อ (ความยาก อำนาจจำแนก) มีการวิเคราะห์ค่าความเที่ยง ความตรง และเลือกใช้วิธีการที่เหมาะสม

22. มีการทดลองใช้เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 (try out)
- 0 หมายถึง ไม่มีการทดลองใช้เครื่องมือวัด
 - 1 หมายถึง มีการทดลองใช้เครื่องมือวัดแต่ไม่ระบุวัตถุประสงค์ของการทดลองใช้
 - 2 หมายถึง มีการทดลองใช้เครื่องมือวัดจำนวน 1 ครั้ง และระบุวัตถุประสงค์ของการทดลองใช้
 - 3 หมายถึง มีการทดลองใช้เครื่องมือวัดมากกว่า 1 ครั้ง แต่ไม่ระบุวัตถุประสงค์ของการทดลองใช้
 - 4 หมายถึง มีการทดลองใช้เครื่องมือวัดมากกว่า 1 ครั้งและระบุวัตถุประสงค์ของการทดลองใช้
23. ตัวอย่างที่ใช้ในการทดลองเครื่องมือวัดสมรรถนะในศตวรรษที่ 21 มีความเหมาะสม
- 0 หมายถึง ไม่มีการทดลองใช้เครื่องมือวัด
 - 1 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดเป็นส่วนหนึ่งของตัวอย่างที่จะใช้จริง
 - 2 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดสุ่มมาจากประชากรกลุ่มเดียวกัน แต่ไม่ได้ระบุจำนวนของตัวอย่าง
 - 3 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดสุ่มมาจากประชากรกลุ่มเดียวกัน แต่จำนวนที่ใช้ไม่พอเพียง/พอเหมาะ
 - 4 หมายถึง ตัวอย่างที่ทดลองใช้เครื่องมือวัดสุ่มมาจากประชากรกลุ่มเดียวกัน และจำนวนที่ใช้พอเพียง/พอเหมาะ
24. เครื่องมือวัดสมรรถนะในศตวรรษที่ 21 ในภาพรวมมีคุณภาพเหมาะสม
- 0 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับต่ำมาก (ค่าความเที่ยงไม่ถึง 0.7 และไม่รายงานค่าความยาก อำนาจจำแนก/ไม่เป็นไปตามเกณฑ์)
 - 1 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับต่ำ (ค่าความเที่ยงไม่ถึง 0.7 ค่าความยาก และอำนาจจำแนกเป็นไปตามเกณฑ์)
 - 2 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับปานกลาง (ค่าความเที่ยง 0.7 ขึ้นไป ค่าความยาก และอำนาจจำแนกเป็นไปตามเกณฑ์)
 - 3 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับดี (ค่าความเที่ยง 0.8 ขึ้นไป ค่าความยาก และอำนาจจำแนกเป็นไปตามเกณฑ์)
 - 4 หมายถึง คุณภาพของเครื่องมือวัดในภาพรวมอยู่ในระดับดีมาก (ค่าความเที่ยง 0.9 ขึ้นไป ค่าความยาก และอำนาจจำแนก เป็นไปตามเกณฑ์)
25. กระบวนการเก็บรวบรวมข้อมูลมีความชัดเจนและเหมาะสม
- 0 หมายถึง ไม่มีการระบุถึงกระบวนการเก็บรวบรวมข้อมูล
 - 1 หมายถึง ระบุถึงกระบวนการเก็บรวบรวมข้อมูลบางส่วน ไม่ระบุเป็นขั้นตอน อ่านแล้วสับสน
 - 2 หมายถึง ระบุถึงกระบวนการเก็บรวบรวมข้อมูลเป็นขั้นตอน แต่วิธีการให้รายละเอียดไม่ชัดเจน
 - 3 หมายถึง ระบุถึงกระบวนการเก็บรวบรวมข้อมูลเป็นขั้นตอน วิธีการเก็บข้อมูลเหมาะสม ข้อความไม่กระชับ เยิ่นเย้อ อ่านแล้วเข้าใจยาก
 - 4 หมายถึง ระบุถึงกระบวนการเก็บรวบรวมข้อมูลเป็นขั้นตอน วิธีการเก็บข้อมูลเหมาะสม ใช้ข้อความกระชับ อ่านแล้วเข้าใจง่าย
26. สถิติที่ใช้ในการวิเคราะห์ข้อมูลมีความถูกต้องเหมาะสมกับลักษณะของข้อมูล
- 0 หมายถึง สถิติที่ใช้ในการวิเคราะห์ข้อมูลไม่สอดคล้องกับวัตถุประสงค์/ปัญหาวิจัย
 - 1 หมายถึง สถิติที่ใช้ในการวิเคราะห์ข้อมูลสอดคล้องกับวัตถุประสงค์/ปัญหาวิจัยบางส่วน
 - 2 หมายถึง สถิติที่ใช้ในการวิเคราะห์ข้อมูลสอดคล้องกับวัตถุประสงค์/ปัญหาวิจัย และระดับข้อมูลที่นำมาวิเคราะห์
 - 3 หมายถึง สถิติที่ใช้ในการวิเคราะห์ข้อมูลสอดคล้องกับวัตถุประสงค์/ปัญหาวิจัย และระดับข้อมูลที่นำมาวิเคราะห์ มีการตรวจสอบข้อตกลงเบื้องต้นที่ใช้
 - 4 หมายถึง สถิติที่ใช้ในการวิเคราะห์ข้อมูลสอดคล้องกับวัตถุประสงค์/ปัญหาวิจัย และระดับข้อมูลที่นำมาวิเคราะห์ มีการตรวจสอบข้อตกลงเบื้องต้นที่ใช้ และปรับแก้ข้อมูลให้สามารถนำมาใช้ในการทดสอบสมมติฐานได้
27. ลักษณะการนำเสนอการแปลผลการวิเคราะห์ข้อมูล

- 0 หมายถึง ไม่มีการนำเสนอผลการวิเคราะห์ข้อมูล
- 1 หมายถึง มีการนำเสนอผลการวิเคราะห์ข้อมูลที่ควรจะนำเสนอ แต่ไม่ครบถ้วน
- 2 หมายถึง มีการนำเสนอผลการวิเคราะห์ข้อมูลที่ควรจะนำเสนออย่างครบถ้วน แต่วิธีนำเสนอไม่เหมาะสม เช่น แสดงเป็นข้อความ อ่านเข้าใจยาก
- 3 หมายถึง มีการนำเสนอผลการวิเคราะห์ข้อมูลที่ควรจะนำเสนออย่างครบถ้วน วิธีนำเสนอเหมาะสม เช่น นำเสนอในรูปของตาราง แต่ไม่มีการอธิบายผลการวิเคราะห์
- 4 หมายถึง มีการนำเสนอผลการวิเคราะห์ข้อมูลที่ควรจะนำเสนออย่างครบถ้วน วิธีนำเสนอเหมาะสม เช่น นำเสนอในรูปของตาราง มีการอธิบายผลการวิเคราะห์
28. การแปลผลและสรุปผลการวิเคราะห์ข้อมูลมีความถูกต้องชัดเจน
- 0 หมายถึง ไม่มีการแปลความหมายหรือสรุปผลการวิเคราะห์ข้อมูล
- 1 หมายถึง มีการแปลความหมายหรือสรุปผลการวิเคราะห์ข้อมูลเป็นบางส่วนเท่านั้น และไม่ถูกต้องตามผลวิเคราะห์ที่ได้
- 2 หมายถึง มีการแปลความหมายหรือสรุปผลการวิเคราะห์ข้อมูลเป็นบางส่วนเท่านั้น และถูกต้องสอดคล้องกับผลวิเคราะห์ที่ได้
- 3 หมายถึง มีการแปลความหมายหรือสรุปผลการวิเคราะห์ข้อมูลครบทุกส่วน แต่บางส่วนไม่ถูกต้องไม่สอดคล้องกับผลวิเคราะห์ที่ได้
- 4 หมายถึง มีการแปลความหมายหรือสรุปผลการวิเคราะห์ข้อมูลครบทุกส่วน ถูกต้องและสอดคล้องกับผลวิเคราะห์ที่ได้
29. สรุปผลการวิจัยอย่างถูกต้อง
- 0 หมายถึง ไม่มีการสรุปผลการวิจัย
- 1 หมายถึง สรุปผลวิจัยไม่สอดคล้องกับวัตถุประสงค์การวิจัย หรือข้อมูลที่ได้จากการค้นพบ
- 2 หมายถึง สรุปผลวิจัยสอดคล้องกับวัตถุประสงค์การวิจัยบางส่วน แต่ไม่ครอบคลุมทุกหัวข้อ
- 3 หมายถึง สรุปผลวิจัยถูกต้องสอดคล้องกับวัตถุประสงค์การวิจัย ตอบปัญหาการวิจัยครอบคลุมทุกหัวข้อ แต่ข้อความไม่กระชับ อ่านเข้าใจยาก ไม่เข้าใจถึงประเด็นที่ชัดเจนที่ต้องการสรุปในหัวข้อนั้น ๆ
- 4 หมายถึง สรุปผลวิจัยถูกต้องสอดคล้องกับวัตถุประสงค์การวิจัย ตอบปัญหาวิจัยครอบคลุมทุกหัวข้อ ใช้ข้อความกระชับ อ่านเข้าใจง่าย อ่านแล้วเข้าใจถึงประเด็นที่ต้องการสรุปในหัวข้อนั้น ๆ อย่างชัดเจน
30. มีการอภิปรายผลการวิจัยสอดคล้องกับผลงานวิจัยในอดีต
- 0 หมายถึง ไม่มีการอภิปรายผลการวิจัยที่ได้กับผลงานวิจัยในอดีต
- 1 หมายถึง มีการอภิปรายผลการวิจัยแต่เป็นข้อคิดเห็นส่วนตัวของผู้วิจัย ไม่ครอบคลุมทุกประเด็นปัญหาวิจัย
- 2 หมายถึง มีการอภิปรายผลการวิจัยที่ได้กับผลงานวิจัยในอดีตอย่างไม่สมเหตุผล ไม่ครอบคลุมทุกประเด็นปัญหาวิจัย
- 3 หมายถึง มีการอภิปรายผลการวิจัยที่ได้กับผลงานวิจัยในอดีตอย่างสมเหตุผล แต่ไม่ครอบคลุมทุกประเด็นปัญหาวิจัย
- 4 หมายถึง มีการอภิปรายผลการวิจัยที่ได้กับผลงานวิจัยในอดีตอย่างสมเหตุผล ครอบคลุมทุกประเด็นปัญหาวิจัย
31. ข้อเสนอแนะมีความชัดเจนและเป็นประโยชน์
- 0 หมายถึง ไม่มีการเขียนข้อเสนอแนะ
- 1 หมายถึง มีข้อเสนอแนะแต่ไม่เกี่ยวข้องกับเรื่องที่วิจัย
- 2 หมายถึง มีข้อเสนอแนะถึงการนำผลการวิจัยไปใช้ไม่สมเหตุผล เป็นข้อคิดเห็นส่วนตัวของผู้วิจัยไม่ได้มาจากการวิจัย
- 3 หมายถึง มีข้อเสนอแนะถึงการนำผลการวิจัยไปใช้มาจากผลการวิจัย มีเหตุผลรองรับเพียงพอ
- 4 หมายถึง มีข้อเสนอแนะถึงการนำผลการวิจัยไปใช้มาจากผลการวิจัย มีเหตุผลรองรับเพียงพอ สามารถมองเห็นแนวทางนำไปใช้ประโยชน์

32. เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางปฏิบัติ
- 0 หมายถึง ผลการวิจัยไม่สร้างองค์ความรู้ใหม่
 - 1 หมายถึง ผลการวิจัยสร้างองค์ความรู้ใหม่ สามารถนำไปใช้ประโยชน์ได้เฉพาะในระดับกลุ่มบุคคล
 - 2 หมายถึง ผลการวิจัยสร้างองค์ความรู้ใหม่ สามารถนำไปใช้ประโยชน์ได้เฉพาะในระดับหน่วยงาน
 - 3 หมายถึง ผลการวิจัยสร้างองค์ความรู้ใหม่ สามารถนำไปใช้ประโยชน์ได้เฉพาะในระดับจังหวัด
 - 4 หมายถึง ผลการวิจัยสร้างองค์ความรู้ใหม่ สามารถนำไปใช้ประโยชน์ได้เฉพาะในระดับประเทศ
33. เป็นงานวิจัยที่นำไปใช้ประโยชน์ได้ในทางวิชาการ
- 0 หมายถึง มีประโยชน์ในด้านวิชาการเฉพาะตัวผู้วิจัย
 - 1 หมายถึง มีประโยชน์ในด้านวิชาการระดับการประยุกต์ทฤษฎีเพื่อนำไปใช้
 - 2 หมายถึง มีประโยชน์ในด้านวิชาการระดับที่ได้องค์ความรู้ใหม่
 - 3 หมายถึง มีประโยชน์ในด้านวิชาการระดับที่ได้องค์ความรู้ใหม่ ปรับปรุงทฤษฎีและแนวคิด
 - 4 หมายถึง มีประโยชน์ในด้านวิชาการระดับที่ได้องค์ความรู้ใหม่ ปรับปรุงทฤษฎีและแนวคิด รวมไปถึงการประยุกต์ทฤษฎีเพื่อนำไปใช้
34. รูปแบบรายงานถูกต้องตามหลักวิชา (5 ส่วน คือ บทนำ รายงานเอกสารที่เกี่ยวข้อง วิธีการวิจัย ผลการวิเคราะห์ข้อมูล และสรุปและอภิปรายผล)
- 0 หมายถึง รูปแบบรายงานถูกต้องเพียง 1 ส่วน
 - 1 หมายถึง รูปแบบรายงานถูกต้อง 2 ส่วน
 - 2 หมายถึง รูปแบบรายงานถูกต้อง 3 ส่วน
 - 3 หมายถึง รูปแบบรายงานถูกต้อง 4 ส่วน
 - 4 หมายถึง รูปแบบรายงานถูกต้อง 5 ส่วน

แบบบันทึกคุณลักษณะงานวิจัย

ชื่อผู้วิจัย.....

ชื่องานวิจัย.....

- | | | | |
|---|---|--------------------------|--------------------------|
| 1. รหัสงานวิจัย | | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. ปีที่พิมพ์งานวิจัย | | <input type="checkbox"/> | <input type="checkbox"/> |
| 3. ประเภทสถาบันที่ผลิตงานวิจัย | | <input type="checkbox"/> | |
| 4. สถาบันที่ผลิตงานวิจัย | | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. สาขาที่ผลิตงานวิจัย | | <input type="checkbox"/> | |
| 6. ประเภทงานวิจัย | | <input type="checkbox"/> | |
| 7. เพศของผู้วิจัย | | <input type="checkbox"/> | |
| 8. วัตถุประสงค์ของการวิจัย | | | |
| (1) พัฒนาตัวบ่งชี้/องค์ประกอบ | (5) พัฒนาเกณฑ์การประเมิน/เกณฑ์การให้คะแนน | <input type="checkbox"/> | <input type="checkbox"/> |
| (2) สร้าง/พัฒนาเครื่องมือวัด | (6) ศึกษาระดับ/เปรียบเทียบคุณลักษณะที่วัดกับตัวแปรอิสระ | <input type="checkbox"/> | <input type="checkbox"/> |
| (3) ตรวจสอบคุณภาพเครื่องมือวัด | (7) ทดสอบความไม่แปรเปลี่ยน | <input type="checkbox"/> | <input type="checkbox"/> |
| (4) สร้างปกติวิสัย (norms) | (8) สร้างคู่มือการใช้เครื่องมือวัด | <input type="checkbox"/> | <input type="checkbox"/> |
| 9. เครื่องมือวัดที่พัฒนาขึ้น | | <input type="checkbox"/> | <input type="checkbox"/> |
| 10. แนวคิด/ทฤษฎีหลัก | | <input type="checkbox"/> | <input type="checkbox"/> |
| 11. องค์ประกอบที่มุ่งวัด | | | |
| (1) จิตสำนึกต่อโลก | (7) ทักษะด้านการทำงานเป็นทีม ความร่วมมือ | <input type="checkbox"/> | <input type="checkbox"/> |
| (2) ความรู้เรื่องพลเมือง | (8) ภาวะผู้นำ | <input type="checkbox"/> | <input type="checkbox"/> |
| (3) การคิดอย่างมีวิจารณ์ญาณ | (9) ความรับผิดชอบ | <input type="checkbox"/> | <input type="checkbox"/> |
| (4) การแก้ปัญหา | (10) การรู้เท่าทันสื่อและสารสนเทศ | <input type="checkbox"/> | <input type="checkbox"/> |
| (5) การสร้างสรรค์และนวัตกรรม | (11) การรู้เท่าทันเทคโนโลยี | <input type="checkbox"/> | <input type="checkbox"/> |
| (6) การสื่อสาร | | <input type="checkbox"/> | |
| 12. เนื้อหาที่เกี่ยวข้อง (กลุ่มสาระการเรียนรู้) | | <input type="checkbox"/> | |
| 13. บริบทที่ศึกษา | | | |
| 13.1 ภูมิภาค 13.2 แหล่งที่มาของตัวอย่าง | | <input type="checkbox"/> | <input type="checkbox"/> |
| 14. ระดับชั้นของตัวอย่าง | | <input type="checkbox"/> | <input type="checkbox"/> |
| 15. การกำหนดขนาดของตัวอย่าง | | <input type="checkbox"/> | <input type="checkbox"/> |
| 16. ขนาดของตัวอย่างในการวิจัย | | <input type="checkbox"/> | <input type="checkbox"/> |
| 17. วิธีการเลือกตัวอย่าง | | <input type="checkbox"/> | |
| 18. การศึกษาแนวคิดการสร้างเครื่องมือวัด | | <input type="checkbox"/> | |
| 19. การสร้างเครื่องมือวัด | | <input type="checkbox"/> | |
| 20. การกำหนดนิยามเชิงปฏิบัติการ | | <input type="checkbox"/> | |

21. การระบุขั้นตอนในการพัฒนาเครื่องมือวัด
22. การสร้างตารางกำหนดโครงสร้างของเครื่องมือวัด
23. ประเภทของเครื่องมือวัด
- | | | | |
|--|--------------------------------|--------------------------|--------------------------|
| (1) แบบสอบถามหลายตัวเลือก | (6) แบบตรวจสอบรายการ | <input type="checkbox"/> | <input type="checkbox"/> |
| (2) แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า | (7) มาตรฐานประมาณค่า | <input type="checkbox"/> | <input type="checkbox"/> |
| (3) แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า | (8) แบบสังเกต | <input type="checkbox"/> | <input type="checkbox"/> |
| (4) แบบสอบอัตนัย | (9) แบบสอบอัตนัยประยุกต์ (MEQ) | <input type="checkbox"/> | <input type="checkbox"/> |
| (5) แบบสอบภาคปฏิบัติ | (10) อื่น ๆ | <input type="checkbox"/> | <input type="checkbox"/> |
24. จำนวนตัวเลือก/เกณฑ์
25. จำนวนผู้เชี่ยวชาญในการตรวจสอบความตรงตามเนื้อหา
- 25.1 จำนวนผู้เชี่ยวชาญด้านการวัดผล
- 25.2 จำนวนผู้เชี่ยวชาญด้านเนื้อหา
26. จำนวนครั้งที่ใช้ในการทดลอง (เพื่อหาคุณภาพเครื่องมือ นับรวมทั้งหมด)
27. จำนวนข้อที่สร้างก่อนการทดลองใช้ (รวมทั้งหมด)
28. จำนวนข้อที่ใช้จริง (รวมทั้งหมด)
29. ทฤษฎีที่ใช้ในการวิเคราะห์หาคุณภาพเครื่องมือวัด
30. การวิเคราะห์ข้อมูลเพื่อตอบวัตถุประสงค์การวิจัย
- | | | | |
|------------------------|-----------------------------------|--------------------------|--------------------------|
| (1) content analysis | (10) ANCOVA | <input type="checkbox"/> | <input type="checkbox"/> |
| (2) descriptive stat | (11) MANOVA | <input type="checkbox"/> | <input type="checkbox"/> |
| (3) independent t-test | (12) exploratory factor analysis | <input type="checkbox"/> | <input type="checkbox"/> |
| (4) dependent t-test | (13) confirmatory factor analysis | <input type="checkbox"/> | <input type="checkbox"/> |
| (5) chi-square | (14) ทดสอบความไม่แปรเปลี่ยน | <input type="checkbox"/> | <input type="checkbox"/> |
| (6) correlation | (15) ค่าความเที่ยง | <input type="checkbox"/> | <input type="checkbox"/> |
| (7) regression | (16) ค่าความตรง | <input type="checkbox"/> | <input type="checkbox"/> |
| (8) one-way ANOVA | (17) ค่าความยาก | <input type="checkbox"/> | <input type="checkbox"/> |
| (9) two-way ANOVA | (18) ค่าอำนาจจำแนก | <input type="checkbox"/> | <input type="checkbox"/> |
31. การสร้างเกณฑ์ปกติ
- 31.1 ประเภทของเกณฑ์ปกติ
- 31.2 วิธีที่ใช้ในการสร้างเกณฑ์ปกติ
32. คู่มือการใช้เครื่องมือวัด
33. คะแนนประเมินคุณภาพงานวิจัย

34. การทดลองใช้	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3	ครั้งที่ 4	ครั้งที่ 5
34.1 ขนาดของตัวอย่าง	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
34.2 วิธีการเลือกตัวอย่างในการทดลองใช้	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34.3 วัตถุประสงค์ของการทดลองใช้					
(1) ตรวจสอบค่าความเที่ยง	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(2) ตรวจสอบค่าความตรง	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(3) ตรวจสอบความยาก	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(4) ตรวจสอบอำนาจจำแนก	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(5) ตรวจสอบภาษา	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(6) ตรวจสอบเวลาที่ใช้	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(7) ตรวจสอบตัวเลือก	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(8) ตรวจสอบความครอบคลุมของเนื้อหา	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(9) ตรวจสอบคำชี้แจง/วิธีดำเนินการสอบ	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(10) ตรวจสอบความไม่แปรเปลี่ยน	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
34.4 ดัชนีเสถียรคุณภาพ (การระบุเกณฑ์พิจารณา)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(1) ประเภทความเที่ยง	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
(2) ประเภทความตรง	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
(3) วิธีการหาค่าความยาก	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>
(4) วิธีการหาค่าอำนาจจำแนก	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>

คู่มือสำหรับบันทึกคุณลักษณะงานวิจัย (code book)

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
1. รหัสงานวิจัย	RID	01-99
2. ปีที่พิมพ์งานวิจัย	YEAR	เลขท้าย 2 ตัวของปี พ.ศ.
3. ประเภทสถาบันที่ผลิตงานวิจัย	TYPEUNIV	1 = มหาวิทยาลัยของรัฐ 2 = มหาวิทยาลัยในกำกับของรัฐ 3 = มหาวิทยาลัยราชภัฏ 4 = มหาวิทยาลัยเอกชน 5 = สถาบันอื่น ๆ (รายงานวิจัย)
4. สถาบันที่ผลิตงานวิจัย	UNIV	01 = จุฬาลงกรณ์มหาวิทยาลัย 02 = มหาวิทยาลัยเกษตรศาสตร์ 03 = มหาวิทยาลัยเชียงใหม่ 04 = มหาวิทยาลัยขอนแก่น 05 = มหาวิทยาลัยทักษิณ 06 = มหาวิทยาลัยนเรศวร 07 = มหาวิทยาลัยบูรพา 08 = มหาวิทยาลัยมหาสารคาม 09 = มหาวิทยาลัยศรีนครินทรวิโรฒ 10 = มหาวิทยาลัยสงขลานครินทร์ 11 = มหาวิทยาลัยสุโขทัยธรรมาธิราช 12 = มหาวิทยาลัยราชภัฏกาญจนบุรี 13 = มหาวิทยาลัยราชภัฏนครราชสีมา 14 = มหาวิทยาลัยราชภัฏเพชรบูรณ์ 15 = มหาวิทยาลัยราชภัฏมหาสารคาม 16 = มหาวิทยาลัยราชภัฏเลย 17 = มหาวิทยาลัยราชภัฏอุบลราชธานี 18 = สถาบันทดสอบทางการศึกษาแห่งชาติ (องค์การมหาชน) 19 = สถาบันอื่น (รายงานวิจัย)
5. สาขาที่ผลิตงานวิจัย	MAJOR	1 = การวัดผลการศึกษา, การวัดและประเมินผลการศึกษา 2 = วิธีวิทยาการวิจัยการศึกษา 3 = วิจัยและประเมินผลการศึกษา, การวัดผลและการวิจัยการศึกษา, วัดผลและสถิติการศึกษา. 4 = เทคโนโลยีการวัดทางการศึกษา. 5 = จิตวิทยาการแนะแนว 6 = สาขาอื่น ๆ 7 = แหล่งอื่น (รายงานวิจัย)
6. ประเภทงานวิจัย	TYPES	1 = วิทยานิพนธ์ ป.โท 2 = วิทยานิพนธ์ ป.เอก

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
		3 = รายงานวิจัย
7. เพศของผู้วิจัย	SEX	1 = ชาย 2 = หญิง 3 = ชายและหญิง
8. วัตถุประสงค์ของการวิจัย (1) พัฒนาตัวบ่งชี้/องค์ประกอบ (2) สร้าง/พัฒนาเครื่องมือวัด (3) ตรวจสอบคุณภาพเครื่องมือวัด (4) สร้างปกติวิสัย (norms) (5) พัฒนาเกณฑ์การประเมิน/เกณฑ์ การให้คะแนน (6) ศึกษาระดับ/เปรียบเทียบ คุณลักษณะที่วัดกับตัวแปรอิสระ (7) ทดสอบความไม่แปรเปลี่ยน (8) สร้างคู่มือการใช้เครื่องมือวัด	OBJECT	0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่
9. เครื่องมือวัดที่พัฒนาขึ้น	TEST	01 = เครื่องมือวัดทักษะแห่งศตวรรษที่ 21 02 = เครื่องมือวัดความตระหนักรู้ต่อโลก 03 = เครื่องมือวัดคุณลักษณะอันพึงประสงค์ (ความรับผิดชอบ, คุณธรรม จริยธรรม และ ค่านิยมที่พึงประสงค์, ความเป็นพลเมืองดี) 04 = เครื่องมือวัดความมีวินัยในตนเอง 05 = เครื่องมือวัดทักษะชีวิต/ทักษะชีวิตและ อาชีพ 06 = เครื่องมือวัดทักษะการคิด/ทักษะการคิด ขั้นสูง 07 = เครื่องมือวัดความสามารถการคิดอย่างมี วิจารณญาณ 08 = เครื่องมือวัดทักษะการแก้ปัญหา (คณิตศาสตร์, วิทยาศาสตร์) 09 = เครื่องมือวัดความคิดสร้างสรรค์ 10 = เครื่องมือวัดทักษะกระบวนการทาง คณิตศาสตร์ 11 = เครื่องมือวัดจิตวิทยาศาสตร์ 12 = เครื่องมือวัดความสามารถในการอ่าน/ เขียน/อ่าน คิดวิเคราะห์ และเขียนข้อความ 13 = เครื่องมือวัดความสามารถทางด้าน เทคโนโลยีสารสนเทศ (จริยธรรมในการใช้ เทคโนโลยีสารสนเทศ, ความไว้วางใจระหว่าง บุคคลในเครือข่ายสังคมออนไลน์, พฤติกรรม การติดอินเทอร์เน็ต) 14 = เครื่องมือวัดทักษะการรู้เท่าทันสื่อ/การ รู้เท่าทันสื่อและสารสนเทศ/การรู้ดิจิทัล

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
(7) ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (8) ภาวะผู้นำ (9) ความรับผิดชอบ (10) การรู้เท่าทันสื่อและสารสนเทศ (11) การรู้เท่าทันเทคโนโลยี		0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่
12. เนื้อหาที่เกี่ยวข้อง (กลุ่มสาระการเรียนรู้)	CONTENT	1 = ภาษาไทย 2 = คณิตศาสตร์ 3 = วิทยาศาสตร์ 4 = สังคมศึกษา ศาสนา และวัฒนธรรม 5 = สุขศึกษาและพลศึกษา 6 = ศิลปะ (ทัศนศิลป์ ดนตรี และนาฏศิลป์) 7 = การงานอาชีพและเทคโนโลยี 8 = ภาษาต่างประเทศ 9 = ไม่สังกัดกลุ่มสาระ (content free)
13. บริบทที่ศึกษา		
13.1 ภูมิภาค	DISTRIC	1 = กทม. 2 = ภาคเหนือ 3 = ภาคตะวันออกเฉียงเหนือ 4 = ภาคกลาง 5 = ภาคใต้ 6 = มากกว่า 1 ภูมิภาค
13.2 แหล่งที่มาของตัวอย่าง	SOURCE	1 = สำนักงานคณะกรรมการการศึกษาขั้นพื้นฐาน (สพฐ.) 2 = สำนักงานคณะกรรมการส่งเสริมการศึกษาเอกชน 3 = สังกัดองค์กรปกครองส่วนท้องถิ่น (อปท.) 4 = สำนักงานการศึกษา กรุงเทพมหานคร (กทม.) 5 = สังกัดคณะกรรมการการอุดมศึกษา (สกอ.) 6 = อื่น ๆ 9 = ไม่ระบุ
14. ระดับชั้นของตัวอย่าง	S_GRADE	1 = ชั้นมัธยมศึกษาปีที่ 1 2 = ชั้นมัธยมศึกษาปีที่ 2 3 = ชั้นมัธยมศึกษาปีที่ 3 4 = ชั้นมัธยมศึกษาปีที่ 4 5 = ชั้นมัธยมศึกษาปีที่ 5 6 = ชั้นมัธยมศึกษาปีที่ 6 7 = ชั้นมัธยมศึกษาตอนต้น (มากกว่า 1 ชั้นเรียน) 8 = ชั้นมัธยมศึกษาตอนปลาย (มากกว่า 1 ชั้นเรียน)

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
		9 = ชั้นมัธยมศึกษาตอนต้นและตอนปลาย 10 = ชั้นประถมศึกษาและมัธยมศึกษา
15. การกำหนดขนาดของตัวอย่าง	SAMPSIZE	01= สูตรของ Krejcie & Morgan 1970 02= สูตรของ Yamane 1973 03= สูตรของ Darwin 04= สูตรคำนวณของบุญชม ศรีสะอาด 2553 05= เปิดตารางของ ศิริชัย กาญจนวาสี 06= กำหนดขนาดตัวอย่างตามสัดส่วน 07= เกณฑ์ร้อยละของประชากร 08= สูตรกำหนดขนาดด้วยการสุ่มแบบแบ่งชั้น 09= เกณฑ์กำหนดขนาดตัวอย่างสำหรับพัฒนาเครื่องมือ ของ บุญชม ศรีสะอาด 2534 10= โปรแกรม G* Power 11= 10 เท่าของตัวแปร/พารามิเตอร์ 12= 15 เท่าของตัวแปร/พารามิเตอร์ 13= 20 เท่าของตัวแปร/พารามิเตอร์ 14= หลาย ๆ เกณฑ์ 15= อื่น ๆ 99= ไม่ระบุ
16. ขนาดของตัวอย่างในการวิจัย (รวม)	SUMSIZE	0001-9999
17. วิธีการเลือกตัวอย่าง	SAMPLING	1 = ใช้ประชากร/กลุ่มเป้าหมาย 2 = สุ่มอย่างง่าย 3 = สุ่มแบบกลุ่ม 4 = สุ่มแบบแบ่งชั้น 5 = สุ่มแบบสองขั้นตอน 6 = สุ่มแบบหลายขั้นตอน 7 = เลือกอย่างเจาะจง 8 = อื่น ๆ 9 = ไม่ระบุ
18. การศึกษาแนวคิดการสร้างเครื่องมือวัด	LITUR	0 = ไม่มี 1 = มี เฉพาะการวิเคราะห์คุณภาพเครื่องมือ 2 = มี เฉพาะวิธีการสร้างเครื่องมือ 3 = มี ทั้งวิธีการสร้างและการวิเคราะห์คุณภาพ
19. การสร้างเครื่องมือวัด	INSTRU	1 = ผู้วิจัยพัฒนาเครื่องมือวัดขึ้นเอง 2 = ผู้วิจัยใช้เครื่องมือวัดจากงานวิจัยอื่น 3 = ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยไทย 4 = ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยต่างประเทศ 5 = ผู้วิจัยปรับปรุงเครื่องมือวัดจากงานวิจัยไทยและต่างประเทศ

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
20. การกำหนดนิยามเชิงปฏิบัติการ	OPER_DEFI	0 = ไม่มี 1 = มี แต่ไม่นำเสนอ 2 = มี และนำเสนอ
21. การระบุขั้นตอนในการพัฒนา เครื่องมือวัด	STEP_SCALE	0 = ไม่มี 1 = มี
22. การสร้างตารางกำหนดโครงสร้างของ เครื่องมือวัด	TESTBLU	0 = ไม่มี 1 = มี แต่ไม่นำเสนอ 2 = มี และนำเสนอ
23. ประเภทของเครื่องมือวัด (1) แบบสอบหลายตัวเลือก (2) แบบวัดเชิงสถานการณ์ แบบให้ คะแนน 2 ค่า (3) แบบวัดเชิงสถานการณ์ แบบให้ คะแนนมากกว่า 2 ค่า (4) แบบสอบอัตนัย (5) แบบสอบภาคปฏิบัติ (6) แบบตรวจสอบรายการ (7) มาตรฐานประมาณค่า (8) แบบสังเกต (9) แบบสอบอัตนัยประยุกต์ (MEQ) (10) อื่น ๆ	INSTTYPE	0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่
24. จำนวนตัวเลือก/เกณฑ์	CHOICE	1 = ไม่มีตัวเลือก 2 = 2 ตัวเลือก/เกณฑ์ 3 = 3 ตัวเลือก/เกณฑ์ 4 = 4 ตัวเลือก/เกณฑ์ 5 = 5 ตัวเลือก/เกณฑ์ 6 = มากกว่า 5 ตัวเลือก/เกณฑ์ 7 = ใช้มากกว่า 1 ประเภท
25. จำนวนผู้เชี่ยวชาญในการตรวจสอบ ความตรงตามเนื้อหา	NUMEXP	ระบุจำนวนตามจริง 0 = ไม่มีผู้เชี่ยวชาญ 9 = ไม่ระบุ
25.1 จำนวนผู้เชี่ยวชาญด้านการ วัดผล	EVAEXP	ระบุจำนวนตามจริง 0 = ไม่มีผู้เชี่ยวชาญ 9 = ไม่ระบุ
25.2 จำนวนผู้เชี่ยวชาญด้านการ เนื้อหา	SPEEXP	ระบุจำนวนตามจริง 0 = ไม่มีผู้เชี่ยวชาญ 9 = ไม่ระบุ
26. จำนวนครั้งที่ใช้ในการทดลอง	NUMTRY	0 = ไม่มี 1 = 1 ครั้ง 2 = 2 ครั้ง 3 = 3 ครั้ง 4 = 4 ครั้ง 5 = 5 ครั้ง 9 = ไม่มี/ ไม่ระบุ
27. จำนวนข้อที่สร้างก่อนการทดลองใช้ (รวมทั้งหมด)	PREITIEM	001-999 999 = ไม่ระบุ
28. จำนวนข้อที่ใช้จริง (รวมทั้งหมด)	ITIEM	001-999 999 = ไม่ระบุ

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
		6 = วิธีเปอร์เซ็นไทล์และสเตรโนนร่วมกัน 7 = วิธีคะแนนมาตรฐานที่ และจำแนกตาม ระดับชั้นเรียนรวมกัน 9 = ไม่ระบุ
32. คู่มือการใช้เครื่องมือวัด	MANUAL	0 = ไม่มี 1 = มี
33. คะแนนประเมินคุณภาพงานวิจัย	SC_ASSE	0.00-4.00
34. การทดลองใช้ครั้งที่ 1-5	TRYOUT_1-5	0 = ไม่มี 1 = มี
34.1 ขนาดของตัวอย่าง	SUMSIZE_1-5	001-999
34.2 วิธีการเลือกตัวอย่างในการ ทดลองใช้	SAMPLING_T1-5	1 = ใช้ประชากร/กลุ่มเป้าหมาย 2 = สุ่มอย่างง่าย 3 = สุ่มแบบกลุ่ม 4 = สุ่มแบบแบ่งชั้น 5 = สุ่มแบบสองขั้นตอน 6 = สุ่มแบบหลายขั้นตอน 7 = เลือกอย่างเจาะจง 8 = อื่น ๆ 9 = ไม่ระบุ
34.3 วัตถุประสงค์ของการทดลองใช้ (1) ตรวจสอบค่าความเที่ยง (2) ตรวจสอบค่าความตรง (3) ตรวจสอบความยาก (4) ตรวจสอบอำนาจจำแนก (5) ตรวจสอบภาษา (6) ตรวจสอบเวลาที่ใช้ (7) ตรวจสอบตัวเลือก (8) ตรวจสอบความครอบคลุมของ เนื้อหา (9) ตรวจสอบคำชี้แจง/ วิธีดำเนินการสอบ (10) ทดสอบความไม่แปรเปลี่ยน	OBJTRY_1-5	0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่ 0 = ไม่ใช่ 1 = ใช่
30.4 ดัชนีแสดงคุณภาพ (การระบุ เกณฑ์พิจารณา)	INDEX_1-5	0 = ไม่มี 1 = มี
(1) ประเภทความเที่ยง	RELIA_1-5	00 = ไม่มีการตรวจสอบ 01 = test-retest method 02 = measure of equivalence 03 = test-retest with equivalent- forms method 04 = split-half methods 05 = KR20/KR21 06 = Cronbach's alpha method 07 = Lovett

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
		08 = Hoyt 's analysis 09 = Kappa of Kohen 10 = ความสอดคล้องระหว่างผู้ประเมิน 11 = ใช้มากกว่า 1 วิธี 12 = อื่น ๆ 99 = ไม่ระบุ
(2) ประเภทความตรง	VALID_1-5	0 = ไม่มีการตรวจสอบ 1 = ความตรงตามเนื้อหา 2 = ความตรงตามสภาพ 3 = ความตรงเชิงทำนาย 4 = ความตรงตามโครงสร้างด้วยการตัดสิน โดยผู้เชี่ยวชาญ 5 = ความตรงตามโครงสร้างด้วยวิธี known group 6 = ความตรงตามโครงสร้างด้วยการ วิเคราะห์เมทริกซ์พหุลักษณะวิธี 7 = ความตรงตามโครงสร้างด้วยการ วิเคราะห์องค์ประกอบ 8 = ใช้มากกว่า 1 วิธี 9 = อื่น ๆ 99 = ไม่ระบุ
(3) วิธีการหาค่าความยาก	DIFFICA_1-5	0 = ไม่มีการตรวจสอบ 1 = สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ 2 = สูตร Whitney and Sabers 3 = โปรแกรม Evana 4 = โปรแกรม B-index 5 = โปรแกรม TAP 6 = IRT 7 = อื่น ๆ 8 = ใช้มากกว่า 1 วิธี 99 = ไม่ระบุ
(4) วิธีการหาค่าอำนาจจำแนก	DISCRI_1-5	0 = ไม่มีการตรวจสอบ 1 = สูตรสัดส่วนกลุ่มสูง-กลุ่มต่ำ 2 = item-total correlation 3 = สถิติทดสอบที (t-test) 4 = Biserial correlation 5 = Point Biserial correlation 6 = สูตร Whitney and Sabers 7 = โปรแกรม Evana 8 = โปรแกรม B-index 9 = โปรแกรม TAP

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
		10 = IRT 11 = อื่น ๆ 12 = มากกว่า 1 วิธี 99 = ไม่ระบุ

แบบบันทึกแบบบันทึกผลการวิจัย

- | | |
|---|---|
| 1. รหัสงานวิจัย | <input type="checkbox"/> <input type="checkbox"/> |
| 2. เครื่องมือวัดฉบับที่ | <input type="checkbox"/> <input type="checkbox"/> |
| 3. องค์ประกอบที่มุ่งวัด | <input type="checkbox"/> <input type="checkbox"/> |
| 4. จำนวนผู้สอบในแต่ละฉบับ | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 5. ประเภทของเครื่องมือวัด | <input type="checkbox"/> <input type="checkbox"/> |
| 6. ลักษณะวิธีตอบของเครื่องมือวัด | <input type="checkbox"/> |
| 7. จำนวนตัวเลือก/เกณฑ์ | <input type="checkbox"/> |
| 8. จำนวนข้อคำถามในแต่ละฉบับ | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 9. เวลาที่ใช้ในการทดสอบ | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 10. การตรวจให้คะแนน | <input type="checkbox"/> |
| 11. Mean | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 12. SD | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 13. ประเภทความเที่ยง | <input type="checkbox"/> |
| 14. ค่าความเที่ยง | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 15. ประเภทและค่าความตรง (เฉพาะที่รายงานค่าสัมประสิทธิ์สหสัมพันธ์) | |
| (1) ความตรงตามสภาพ | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| (2) ความตรงเชิงทำนาย | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 16. ค่าความยาก (ทั้งฉบับ) | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| 17. ประเภทและค่าอำนาจจำแนก (ทั้งฉบับ) | |
| (1) discrimination power (r) | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| (2) item-total correlation | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |
| (3) ค่าสถิติทดสอบ t | <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> |

คู่มือแบบบันทึกผลการวิจัย

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
1. รหัสงานวิจัย	RID	01-999
2. เครื่องมือวัดฉบับที่	TEST	1-99
3. องค์ประกอบที่มุ่งวัด	COMPONENT	01 = จิตสำนึกต่อโลก 02 = ความรู้เรื่องพลเมือง 03 = การคิดอย่างมีวิจารณญาณ 04 = การแก้ปัญหา 05 = การสร้างสรรค์และนวัตกรรม 06 = การสื่อสาร 07 = ทักษะด้านการทำงานเป็นทีม ความร่วมมือ 08 = ภาวะผู้นำ 09 = ความรับผิดชอบ 10 = การรู้เท่าทันสื่อและสารสนเทศ 11 = การรู้เท่าทันเทคโนโลยี
4. จำนวนผู้สอบในแต่ละฉบับ	SAMPLE	0001-9999
5. ประเภทของเครื่องมือวัด	INSTTYPE	01 = แบบสอบหลายตัวเลือก 02 = แบบวัดเชิงสถานการณ์ แบบให้คะแนน 2 ค่า 03 = แบบวัดเชิงสถานการณ์ แบบให้คะแนนมากกว่า 2 ค่า 04 = แบบสอบอัตนัย 05 = แบบสอบภาคปฏิบัติ 06 = แบบตรวจสอบรายการ 07 = มาตรฐานค่า 08 = แบบสังเกต 09 = แบบสอบอัตนัยประยุกต์ MEQ= 10 = อื่น ๆ
6. ลักษณะวิธีตอบของเครื่องมือวัด	ANSTYPE	1 = แบบเขียนตอบ (paper-pencil) 2 = คอมพิวเตอร์ 3 = ปฏิบัติจริง 4 = ปากเปล่า 5 = อื่น ๆ
7. จำนวนตัวเลือก/เกณฑ์	CHOICE	1 = ไม่มีตัวเลือก 2 = 2 ตัวเลือก/เกณฑ์ 3 = 3 ตัวเลือก/เกณฑ์ 4 = 4 ตัวเลือก/เกณฑ์ 5 = 5 ตัวเลือก/เกณฑ์ 6 = มากกว่า 5 ตัวเลือก/เกณฑ์ 7 = ใช้มากกว่า 1 ประเภท

ตัวแปร	ชื่อตัวแปร (ภาษาอังกฤษ)	รหัสและค่าของตัวแปร
8. จำนวนข้อคำถามในแต่ละฉบับ	ITEM	001-999
9. เวลาที่ใช้ในการทดสอบ	TESTTIME	000-999 (นาที)
10. การตรวจให้คะแนน	SCORE	1 = ให้คะแนนแบบสองค่า 2 = ให้คะแนนมากกว่าสองค่า 3 = ให้คะแนนแบบสองค่า และมากกว่าสองค่า 9 = ไม่ระบุ
11. Mean	MEAN	0.00-999.99
12. SD	SD	0.00-999.99
13. ประเภทความเที่ยง	TYPE_RELI	00 = ไม่มีการตรวจสอบ 01 = test-retest method 02 = measure of equivalence 03 = test-retest with equivalent-forms method 04 = split-half methods 05 = KR20/KR21 06 = Cronbach's alpha method 07 = Lovett 08 = Hoyt 's analysis 09 = Kappa of Kohen 10 = ความสอดคล้องระหว่างผู้ประเมิน
14. ค่าความเที่ยง	RELI	0.000-1.000 0.000 = ไม่ตรวจสอบ/ไม่ระบุ
15. ประเภทและค่าความตรง (เฉพาะที่ รายงานค่าสัมประสิทธิ์สหสัมพันธ์) (1) ความตรงตามสภาพ (2) ความตรงเชิงทำนาย	VALID	0.000-1.000 0.000 = ไม่ตรวจสอบ/ไม่ระบุ 0.000-1.000 0.000 = ไม่ตรวจสอบ/ไม่ระบุ
16. ค่าความยาก (ทั้งฉบับ)	P	0.000-1.000 0.000 = ไม่ตรวจสอบ/ไม่ระบุ
17. ประเภทและค่าอำนาจจำแนก (ทั้งฉบับ) (1) discrimination power (r) (2) item-total correlation (3) ค่าสถิติทดสอบ t	DISCRI	0.000-1.000 0.000 = ไม่ตรวจสอบ/ไม่ระบุ 0.000-1.000 0.000 = ไม่ตรวจสอบ/ไม่ระบุ ระบุตามจริง 0.000 = ไม่ตรวจสอบ/ไม่ระบุ

ภาคผนวก ง

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะ
ของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา

**ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะ
ของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา**

ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดสมรรถนะของผู้เรียนในศตวรรษที่ 21 ระดับมัธยมศึกษา ผู้วิจัยได้นำเสนอจำแนกตามสมรรถนะหลักทั้ง 4 สมรรถนะ และสมรรถนะย่อยรวมทั้งสิ้น 11 สมรรถนะ ดังนี้

- 1. ทักษะและความรู้ในเนื้อหาขั้นสูง**
 - 1.1 จิตสำนึกต่อโลก (global awareness)
 - 1.2 ความรู้เรื่องพลเมือง (civic literacy)
- 2. ทักษะการคิดขั้นสูง**
 - 2.1 การคิดอย่างมีวิจารณญาณ (critical thinking)
 - 2.2 การแก้ปัญหา (solving problems)
 - 2.3 การสร้างสรรค์และนวัตกรรม (creativity and innovation)
- 3. ทักษะต่อตนเองและสังคม**
 - 3.1 การสื่อสาร (communications literacy/skill)
 - 3.2 ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (collaboration, teamwork)
 - 3.3 ภาวะผู้นำ (leadership)
 - 3.4 ความรับผิดชอบ (responsibility)
- 4. ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี**
 - 4.1 การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy)
 - 4.2 การรู้เท่าทันเทคโนโลยี (technology literacy)

ซึ่งมีรายละเอียดผลการสังเคราะห์องค์ความรู้แต่ละสมรรถนะ ดังตาราง ง.1 – ง.11

1. ทักษะและความรู้ในเนื้อหาขั้นสูง

1.1 จิตสำนึกต่อโลก (global awareness)

ตาราง ง.1 ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดจิตสำนึกต่อโลก

ประเด็นการสังเคราะห์	งานวิจัยเรื่องที่			ความถี่ (f)
	4	30	42	
1.นิยาม/ตัวบ่งชี้				
ความตระหนักเกี่ยวกับสาเหตุของการเกิดภาวะโลกร้อน		✓		1
ความตระหนักเกี่ยวกับผลกระทบอันเกิดจากภาวะโลกร้อน		✓		1
การปฏิบัติตนเพื่อช่วยลดภาวะโลกร้อน (การลดการใช้ ภาววัสดุทดแทน และการประหยัดพลังงาน การไม่ทำลายสิ่งแวดล้อม)	✓	✓		2
ความตระหนักในมุมมองที่แตกต่าง			✓	1
ความตระหนักในสภาพปัจจุบันของโลก			✓	1
ความตระหนักในความแตกต่างของวัฒนธรรม			✓	1
ความรู้ในเรื่องพลวัตของโลก			✓	1

ประเด็นการสังเคราะห์	งานวิจัยเรื่องที่			ความถี่ (f)
	4	30	42	
ความตระหนักต่อทางเลือกของมนุษย์			✓	1
2. แนวคิดทฤษฎีที่ใช้				
ความตระหนักต่อโลกตามแนวคิดของ Hanvey (1982)			✓	1
ความตระหนักต่อภาวะโลกร้อนตามแนวคิดของ Stamm at al. (2000)		✓		1
ไม่ระบุ	✓			1
3. ชื่อเครื่องมือหลักในงานวิจัย				
แบบวัดความตระหนักผู้รู้ต่อผลกระทบของสภาวะโลกร้อน	✓	✓		2
แบบวัดความตระหนักต่อโลก			✓	1
4. ประเภทเครื่องมือและการให้คะแนน				
มาตรฐานประมาณค่า 5 ระดับ ให้คะแนน 1, 2, 3, 4, 5	✓		✓	2
แบบสอบถามหลายตัวเลือก แบบวัดเชิงสถานการณ์ 5 ตัวเลือก ให้คะแนน 2 ค่า (0,1)		✓	✓	2

1.2 ความรู้เรื่องพลเมือง (civic literacy)

ตาราง ง.2 ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดความรู้เรื่องพลเมือง

สรุปงานวิจัยของ จุฑามาศ ช่อคง (2556)	
1. นิยาม/ตัวบ่งชี้	1) รักชาติศาสน์กษัตริย์ 2) ซื่อสัตย์สุจริต 3) มีวินัย เคารพกฎหมายและกติกาสังคม 4) ใฝ่เรียนรู้ 5) อยู่อย่างพอเพียง 6) มุ่งมั่นรับผิดชอบในการทำงาน 7) รักความเป็นไทย 8) มีจิตสาธารณะ 9) มีเหตุผลรับฟังความคิดเห็นของผู้อื่น และ 10) เข้าร่วมกิจกรรมทางการเมืองการปกครอง
2. แนวคิดทฤษฎีที่ใช้	สังเคราะห์คุณลักษณะของความเป็นพลเมืองดีจากคุณลักษณะอังกพิงประสงค์ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 และคุณลักษณะพลเมืองของประเทศชาติและสังคมโลก ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน พุทธศักราช 2551 กลุ่มสาระการเรียนรู้สังคมศึกษา ศาสนาและวัฒนธรรม
3. ชื่อเครื่องมือหลักในงานวิจัย	แบบวัดความเป็นพลเมืองดี
4. ประเภทเครื่องมือและการให้คะแนน	แบบวัดเชิงสถานการณ์ ให้คะแนนมากกว่า 2 ค่า (0,1,2,3) แบบ 4 ตัวเลือก

ประเด็นการสังเคราะห์	งานวิจัยสำคัญที่												ความถี่ (f)						
	9	10	13	28	32	35	38	45	55	59	60	61		77	79	82	84	88	92
2. แนวคิดทฤษฎีที่ใช้																			
แนวคิดการคิดอย่างมีวิจารณญาณของ Ennis (1985)				✓								✓					✓		4
แนวคิดการคิดอย่างมีวิจารณญาณของ Watson & Glaser (1973)								✓				✓							2
Partnership for 21st Century Learning														✓					2
สังเคราะห์จากหลายแนวคิด/ทฤษฎี		✓				✓			✓										5
อื่นๆ (WHO, สิริชัย กาญจนวาสี และคณะ 2551, แนวคิดจากแบบวัด CAAP)	✓			✓									✓						4
ไม่ระบุ								✓											1
3. ชื่อเครื่องมือหลักในงานวิจัย																			
แบบวัดทักษะ/ความสามารถ/การคิดอย่างมีวิจารณญาณ		✓		✓				✓					✓			✓			9
แบบวัดความสามารถ/ทักษะการคิด (ขั้นสูง)			✓											✓					3
แบบวัดทักษะชีวิต	✓																		1
เครื่องมือประเมินสมรรถนะสำคัญ					✓														1
แบบวัดคุณลักษณะการเรียนรู้						✓													1
แบบวัดคุณลักษณะของผู้เรียน										✓									1
เครื่องมือวัดความสามารถพิเศษทางคณิตศาสตร์													✓						1
4. ประเภทเครื่องมือและการให้คะแนน																			
แบบวัดเชิงสถานการณ์ ให้คะแนนมากกว่า 2 ค่า แบบ 3 ตัวเลือก	✓																		1
แบบสอบหลายตัวเลือก ให้คะแนนสองค่า (0, 1) แบบ 2,4,5 ตัวเลือก		✓		✓				✓					✓				✓		14
แบบมาตราประมาณค่า 5 ระดับ ให้คะแนน 1,2,3,4,5															✓				2
แบบสอบอัตนัย ให้คะแนน 0, 1, 2, 3																✓			1

3.2 ทักษะด้านการทำงานเป็นทีม ความร่วมมือ (collaboration, teamwork)

ตาราง ง.7 ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการทำงานเป็นทีม ความร่วมมือ

ประเด็นการสังเคราะห์	งานวิจัยเรื่องที่										ความถี่ (f)
	11	18	25	26	44	56	72	82	87	92	
1. นิยาม/ตัวบ่งชี้											
ความเต็มใจที่จะทำงานร่วมกับผู้อื่น	✓		✓		✓			✓			4
รู้จักบทบาทของตนที่ได้รับมอบหมาย	✓			✓			✓		✓		4
ให้ความร่วมมือกับผู้อื่น และมีทักษะการทำงานร่วมกับผู้อื่น	✓	✓				✓				✓	4
ยอมรับความคิดเห็นของผู้อื่น		✓							✓	✓	3
เห็นคุณค่าของการทำงานร่วมกับผู้อื่น			✓	✓			✓				3
การบริหารความขัดแย้ง						✓			✓		2
การแลกเปลี่ยนความคิดเห็นกับบุคคลอื่น		✓							✓		2
ประพฤติและปฏิบัติตามข้อตกลงของกลุ่ม			✓								1
การสร้างสัมพันธ์ที่ดีกับผู้อื่น มีมนุษยสัมพันธ์							✓		✓		2
การวางแผนการทำงานร่วมกัน									✓		1
ความคิดเห็นหรือผลงานเป็นที่ยอมรับของผู้อื่น					✓						1
การร่วมมือรวมพลังโดยการมีผู้นำ						✓					1
การร่วมมือรวมพลังโดยการบริหารการดำเนินงาน						✓					1
2. ชื่อเครื่องมือหลักในงานวิจัย											
มาตรวัดจิตวิทยาศาสตร์	✓	✓	✓	✓	✓		✓				6
แบบวัดการร่วมมือรวมพลัง						✓					1
ชุดเครื่องมือประเมินทักษะการเรียนรู้และนวัตกรรม								✓			1
แบบวัดคุณลักษณะอันพึงประสงค์									✓		1
แนวทางการประเมินทักษะการเรียนรู้และนวัตกรรม										✓	1
3. ประเภทเครื่องมือและการให้คะแนน											
มาตรประมาณค่า 5 ระดับ	✓		✓					✓	✓		4
แบบวัดเชิงสถานการณ์ ให้คะแนน 2 ค่า (0,1) แบบ 4 ตัวเลือก		✓									1
แบบวัดเชิงสถานการณ์ ให้คะแนน 2 ค่า (0,1) แบบ 4-5 ตัวเลือก				✓	✓	✓	✓			✓	5
แบบตรวจสอบรายการ								✓		✓	2
แบบประเมินผลงาน								✓		✓	2

3.3 ภาวะผู้นำ (leadership)

ตาราง ง.8 ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดภาวะผู้นำ

ประเด็นการสังเคราะห์	งานวิจัยเรื่อง					ความถี่ (f)
	39	41	52	56	87	
1. นิยาม/ตัวบ่งชี้						
ความสามารถในการทำให้ผู้อื่นคล้อยตามหรือปฏิบัติตามแนวคิดหรือคำสั่งของตน		✓		✓	✓	3
มีความเห็นอกเห็นใจผู้อื่น ไม่เห็นแก่ตัวมีน้ำใจต่อผู้อื่น คิดถึงประโยชน์ส่วนรวม	✓		✓		✓	3
มีความเสียสละ ไม่คิดร้ายผู้อื่น	✓		✓			2
มีเหตุผลไม่ลำเอียง ปราศจากอคติ	✓		✓			2
ใช้อิทธิพลและอำนาจในทางที่ถูกต้อง				✓		1
เป็นแบบอย่างที่ดี				✓		1
มีความคิดริเริ่มและความรับผิดชอบ					✓	1
สามารถตัดสินใจและแก้ไขปัญหาที่เกิดขึ้นได้อย่างเหมาะสม					✓	1
2. ชื่อเครื่องมือหลักในงานวิจัย						
แบบวัดความมีวินัยในตนเอง	✓		✓			2
แบบวัดทักษะชีวิตและอาชีพในศตวรรษที่ 21		✓				1
แบบวัดการร่วมมือรวมพลัง				✓		1
แบบวัดคุณลักษณะอันพึงประสงค์					✓	1
3. ประเภทเครื่องมือและการให้คะแนน						
มาตรฐานประมาณค่า 5 ระดับ	✓		✓		✓	3
แบบวัดเชิงสถานการณ์ ให้คะแนนมากกว่า 2 ค่า แบบ 4 ตัวเลือก	✓	✓		✓		3

4. ทักษะด้านสารสนเทศ สื่อ และเทคโนโลยี

4.1 การรู้เท่าทันสื่อและสารสนเทศ (media and information literacy)

ตาราง ง.10 ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการรู้เท่าทันสื่อและสารสนเทศ

ประเด็นการสังเคราะห์	งานวิจัยเรื่องที่					ความถี่ (f)
	35	36	40	66	74	
1. นิยาม/ตัวบ่งชี้						
การเข้าถึง (access)	✓	✓	✓	✓	✓	5
การประเมิน (evaluate)	✓	✓	✓	✓	✓	5
การสร้างสรรค์ (create)	✓	✓	✓	✓		4
การมีส่วนร่วม (act)	✓	✓	✓		✓	4
การวิเคราะห์ (analyze)	✓	✓	✓	✓		4
การเข้าใจ (understand)					✓	1
การใช้ (use)				✓		1
จริยธรรม (ethic)					✓	1
2. ชื่อเครื่องมือหลักในงานวิจัย						
แบบวัดคุณลักษณะการเป็นบุคคลแห่งการเรียนรู้	✓					1
แบบวัดทักษะในศตวรรษที่ 21 ด้านการรู้เท่าทันสื่อ		✓				1
แบบวัดทักษะการรู้เท่าทันสื่อในศตวรรษที่ 21			✓			1
มาตรวัดการรู้เท่าทันสื่อและสารสนเทศ				✓		1
แบบวัดสถานการณ์การรู้สื่อดิจิทัล					✓	1
3. ประเภทเครื่องมือและการให้คะแนน						
มาตรประมาณค่า 5 ระดับ	✓	✓				2
แบบวัดเชิงสถานการณ์ ให้คะแนนแบบ 2 ค่า (0,1) แบบ 4 ตัวเลือก		✓	✓			2

4.2 การรู้เท่าทันเทคโนโลยี (technology literacy)

ตาราง ง.11 ผลการสังเคราะห์องค์ความรู้เกี่ยวกับงานวิจัยการพัฒนาเครื่องมือวัดการรู้เท่าทันเทคโนโลยี

ประเด็นการสังเคราะห์	งานวิจัยเรื่องที่							ความถี่ (f)
	7	15	32	35	49	53	70	
1. นิยาม/ตัวบ่งชี้								
การแสดงความสามารถในเครือข่ายสังคมออนไลน์	✓							1
ความซื่อสัตย์สุจริตในการใช้เครือข่ายสังคมออนไลน์	✓							1
ความปรารถนาดีในการใช้เครือข่ายสังคมออนไลน์	✓							1
การใช้อินเทอร์เน็ตและสังคมออนไลน์อย่างเหมาะสม					✓			1
ความรู้ด้านเทคโนโลยีสารสนเทศ		✓	✓			✓		3
ทักษะกระบวนการด้านเทคโนโลยีสารสนเทศ		✓	✓	✓		✓		4
เจตคติในด้านเทคโนโลยีสารสนเทศ		✓	✓			✓	✓	4
ความเป็นส่วนตัว ความถูกต้อง ความเป็นเจ้าของ การเข้าถึงข้อมูล							✓	1
2. ชื่อเครื่องมือหลักในงานวิจัย								
แบบวัดความไว้วางใจระหว่างบุคคลในเครือข่ายสังคมออนไลน์	✓							1
เครื่องมือวัดสมรรถนะผู้เรียนด้านเทคโนโลยีสารสนเทศ		✓				✓		2
เครื่องมือประเมินสมรรถนะสำคัญ			✓					1
แบบวัดคุณลักษณะการเป็นบุคคลแห่งการเรียนรู้				✓				1
แบบวัดพฤติกรรมการติดอินเทอร์เน็ต					✓			1
แบบวัดจริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร							✓	1
3. ประเภทเครื่องมือและการให้คะแนน								
มาตราประมาณค่า 4-6 ระดับ	✓	✓		✓	✓	✓	✓	6
แบบสอบหลายตัวเลือก ให้คะแนน 2 (0,1) แบบ 4 ตัวเลือก		✓						1
แบบสังเกตพฤติกรรม ให้คะแนนตามเกณฑ์รูบริก		✓						1
แบบสอบภาคปฏิบัติ			✓		✓	✓		2